WAVELET-BASED CHARACTERIZATION AND STOCHASTIC MODELLING OF PULSE-LIKE GROUND MOTIONS ON THE TIME-FREQUENCY PLANE

by

Anca Lungu

Dissertation submitted in fulfilment of the requirements for the award of

DOCTOR OF PHILOSOPHY

in

STRUCTURAL ENGINEERING

School of Engineering and Mathematical Sciences

City University London

November 2014

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my supervisor Dr *Agathoklis Giaralis* for his continuous support, patience and encouragement. I was very lucky to find in him a devoted mentor always open to listening and discussing any questions that arose in the course of my studies. His perpetual commitment, selfless time and professionalism were essential for the completion of this thesis.

I would also like to thank the *members of staff of the School of Engineering and Mathematical Sciences Department* for valuable advice and support throughout the teaching activities I have undertaken.

I am grateful to *City University London* for the financial support offered throughout my studies. The partial support of the *School of Engineering and Mathematical Sciences Department* is also gratefully acknowledged.

I am truly thankful to Professor *Pavel Alexa* for advising me and opening the routes towards my doctoral education.

I would like to thank all my actual and former colleagues for their support and friendship during my studies. I am especially grateful to *Eric Cheilletz* for endless conversations which broadened my horizons in so many ways; to *Laurențiu Marian* for joining me for yet another degree, for our valuable conversations and for the incessant help he offered me whenever I asked; to *Paschalis Skrekas* and to *Saeed Javdani* for encouraging and tempering me every time I needed.

I cannot express in words my appreciation towards *Persida Fira-Mladinescu* who was like a mother to me during all these years, *Iulia Muntean* for her unceasing guidance and warm support and *Alexandra Crişan* for keeping me optimistic and brave.

A deep and heartfelt thank you to my sister *Silvia Bashir*, one of my main pillars, for her never ending support and unconditional love which kept me harmonious and equilibrated throughout this journey. I warmly thank my sister's family who embraced me with open hearts and offered me amazing holidays.

I'll forever be grateful to my mother *Felicia Lungu* for being an inspiration on how to give out selflessly, how to grow feet on the ground and head in the clouds and how to live bold, but gracefully. Without her none of this would have been possible. This thesis is dedicated to her memory.

DECLARATION

I grant powers of discretion to the University Librarian to allow this dissertation to be copied in whole or in part without further reference to me. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgements.

ABSTRACT

A novel non-separable non-stationary stochastic model for the representation and simulation of pulse-like earthquake ground motions (PLGMs), capable to accurately represent peak elastic and inelastic structural responses, is proposed in this work. Further, the model is employed for assessing the performance of several time-frequency representation techniques (the harmonic wavelet transform, the Meyer wavelet packets transform, the S-transform and the empirical mode decomposition) in capturing salient features of pulse-like accelerograms.

The significantly higher structural demands posed by PLGMs in comparison with similar intensity pulse-free motions led to comprehensive investigations in order to mitigate the damage experienced in the affected areas, such as those located near seismic faults. In this regard, time-frequency analysis methods are frequently employed for the analysis of signals recorded during these events, due to their adaptability to the specific evolutionary behaviour. Alongside with characterization, stochastic modelling of PLGMs is of interest since it allows for systematic variations of the input parameters in order to enhance the understanding of their influence on the structural behaviour. This is particularly useful since only a limited number of PLGMs are available in the existing earthquake databases.

Accordingly, inspired by the time-frequency distribution of their total energy, a versatile PLGM model is defined as a combination of amplitude-modulated stochastic processes. Each process models the time-varying distribution of the energy for adjacent frequency ranges. Two alternative formulations are proposed for representing the low-frequency content characterizing the pulses. Considering a set of pulses from the literature, numerical results show that the pulse models' parameters can be calibrated to simulate in average the structural impact of these pulses represented using the model herein defined. Further, the capability of the PLGM model to generate elastic and inelastic spectral responses matching a given field recorded accelerogram in the mean sense is illustrated. The applicability of the proposed model to account for near-fault effects to spectrum compatible representations of the seismic action is illustrated by generating a fully stochastic process compatible with the response spectrum of the European aseismic code (EC8). Furthermore, the model can be employed in various applications including generation of accelerograms for nonlinear dynamic analyses of structures, probabilistic seismic demand analyses or as input in stochastic dynamic techniques such as statistical linearization.

Finally, the capability of several time-frequency analysis methods to characterize PLGM accelerograms is evaluated through comparative numerical studies within a novel methodology, namely by considering artificial time-histories as samples of the proposed model. The results highlight the potential of the S-transform to be used for pulse identification/extraction and of the harmonic wavelet transform for record characterization/pulse extraction. Additionally, they confirm that from an engineering perspective the structural natural period is an appropriate and representative parameter for the definition of "pulses". Overall, these analyses shed light into the challenges experienced when attempting to detect the pulse content in the accelerograms, in an effort to inform best practices for PLGMs characterization.

CHAPTER 1 : INTRODUCTION

1.1. MOTIVATION

Signals recorded during pulse-like seismic ground motions (PLGMs) distinguish themselves through the presence of unusually high amplitude, long period oscillations termed as "pulses". It has been observed that such ground motions have a particularly strong impact on relatively flexible structures, whose dominant fundamental period is close to the pulse period. Consequently an important amount of research was devoted especially in the past two decades to understanding and characterizing such earthquakes, with the aim of improving the structural behaviour in the affected areas. A major limitation of these studies is the modest amount of field recorded data available, which includes less than 100 records being classified in the literature as pulse-type. As a result, the development of record-based stochastic models to be used for various applications is an essential matter in pulse-like ground motions related research.

Deterministic pulse-like functions were initially employed for the artificial modelling of such earthquakes, acknowledging however the limited representation obtained this way (Somerville, 1998; Makris & Chang, 2000a). Accordingly, the representation of pulse-like ground motions was revised to include stochastic models for simulating the higher frequency content over-riding the pulses. This update was aiming on one hand to produce more realistic accelerograms either in terms of content or in terms of appearance and structural impact. On the other hand, the purpose was to take into account the inherent variability in the seismic ground motions properties (e.g. Mavroeidis & Papageorgiou, 2003). However, since the pulse properties remained still invariant, in the more recent literature probability distributions are established for the

pulse parameters by considering specific ensembles of records (Bray & Rodriguez-Marek, 2004; Dabaghi et al. 2011; Dickinson & Gavin, 2011).

In this context an alternative stochastic modelling approach inspired by the time-frequency distribution of the record's energy is proposed in this work. It employs a simple but popular representation method to shape the energy distribution of pulse-like ground motions, which has been previously used for various structural dynamics matters (e.g. Conte & Peng, 1997; Spanos & Failla, 2004). The approach consists in superposing several amplitude-modulated random processes, obtaining this way a non-separable, non-stationary random process. Existing stochastic models for the higher frequency content are combined with newly defined ones for pulses. The proposed model aims for simplicity, consistency and flexibility in choosing the level of detailing for the generated processes, while offering a fully stochastic representation of pulse-like ground motions. It is likely that the model will be employed in diverse applications such as performance-based earthquake engineering studies (e.g. Taflanidis & Jia, 2011; Taflanidis, 2011) or statistical linearization-based applications for structural engineering (e.g. Spanos & Giaralis, 2008; Spanos & Kougioutzoglou, 2012).

The capability of the herein defined models to simulate the low-frequency content and the corresponding structural impact is explored. Accordingly, a collection 91 pulses previously extracted from field recorded pulse-like accelerograms are modelled individually and the structural responses obtained through the simulations are compared with those corresponding to the actual pulses. Moreover, a given field recorded accelerogram is represented using the pulse-like ground motion model for illustrative purposes. The potential of the model to incorporate pulse-effects in synthetic accelerograms compatible with design response spectra is also investigated, in the context of the Eurocode 8 provisions. Further, the model is used as a test-bed in an investigation gauging the capability of several time-frequency representation techniques

(three wavelet-based, i.e. the harmonic wavelet, Meyer wavelet packets and S transforms, in addition to the empirical mode decomposition) to identify and isolate the characteristic pulses from pulse-like accelerograms. Two methodologies are developed for this purpose: one has been previously used in the literature for this sort of assessments, while the second one has been developed for the purpose of the herein study in order to exploit the advantages of using controlled-input data.

1.2. THESIS ORGANISATION

This dissertation consists of seven chapters and four appendices, followed by a list of references. The introductory *first chapter* presents the motivation of the objectives of this research work, followed by an outline of the thesis. The *second chapter* provides the necessary mathematical background for the signal processing techniques used throughout this work. The advantages and limitations of wavelet-based techniques, together with the main types of wavelet-transform are presented. The generalized harmonic wavelet transform, the Meyer wavelet packets transform, the S-transform and the empirical mode decomposition are revised.

In the *third chapter* a review of the pulse-like ground motions topic is provided. The distinctive phenomenological features characterizing pulse-like records are exemplified and the physical conditions likely to cause them are presented. Further, the parameters used for the quantification of their characteristics and the pulse models currently existing in the literature are summarized. Finally, the approaches used for the identification and extraction of pulses from the recorded time-histories are reviewed; the currently employed models for modelling and simulating such ground motions are also presented.

In the *fourth chapter* a fully non-stationary stochastic model for representing pulse-like ground motions is proposed. Two novel stochastic approaches for modelling pulses are

presented and the model employed in this work for the higher frequency content representation is described. In the *fifth chapter* the pulse models are fitted to a given database of pulses. Further, the pulse-like ground motion model is calibrated against a given accelerogram in order to portray its potential in simulating realistic structural elastic and inelastic responses. Additionally, its applicability for incorporating pulse-like effects in code-compatible accelerograms is demonstrated.

In the *sixth chapter* the techniques used for the characterization of pulse-like records and the identification/extraction of pulses are assessed. In the first part the potential of the generalized harmonic wavelet transform, the Meyer-wavelet packets transform and the S-transform to estimate the underlying energy distribution of pulse-like processes is evaluated. In the latter part, the same techniques, together with the empirical mode decomposition are evaluated by employing them for pulse extraction from artificial accelerograms. Finally, pulses are extracted from several field recorded accelerograms using the approach proposed herein and their quality is assessed in terms of structural responses.

Chapter seven summarizes the main findings presented in this dissertation, acknowledges the limitations of this work and suggests potential future developments.

CHAPTER 2 : TIME-FREQUENCY ANALYSIS TECHNIQUES

2.1. PRELIMINARY REMARKS

The analysis and the characterization of ground motion signals recorded during earthquake events offer valuable information on the seismic source mechanisms and on the recorded structural impact. The amplitude of the energy content varies in both time and frequency, characteristic also known as non-stationarity. Consequently, such signals call for analysis tools able to unfold their energy on the time-frequency plane. In this chapter, several such time-frequency representation (TFR) techniques are briefly reviewed. In particular, the Short Time Fourier Transform, which overcomes some of the Fourier Transform limitations when dealing with non-stationary signals, is revised. Next, the Wavelet Transform is introduced as a technique with superior adaptability to the signal's content. After discussing the properties and characteristics of the wavelet functions, three wavelet-based techniques used throughout this work are presented: the generalized harmonic wavelet transform, the Meyer wavelet packets transform, and the S-transform. Further, an adaptive non-parametric signal processing technique, the empirical mode decomposition, is also discussed.

2.2. THE FOURIER TRANSFORM

The frequency domain representation of a finite energy signal *f (t)* is given by the equation

$$\hat{F}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \tag{2.1}$$

This conversion from the time-domain to the frequency domain representation is known as the Fourier transform (FT) and is based on the agreement that any periodic function of time can be expressed as an infinite summation of harmonics, with specific amplitude, frequency and phase (Newland, 1984). The frequency content of a signal is thus determined by carrying out its convolution with each decomposing harmonic. High values of the coefficients indicate the presence of the harmonic's frequency in the signal. The time domain representation can be recovered by taking the inverse FT given by

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{F}(\omega) e^{i\omega t} d\omega \tag{2.2}$$

According to the Parseval's theorem the energy of the signal is conserved throughout the transform:

$$En = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\infty}^{+\infty} \left|\hat{F}(\omega)\right|^2 d\omega < \infty \tag{2.3}$$

By treating $|f(t)|^2$ as a probability density function, information on the localization of the signal in time and its duration can be obtained by estimating the time average $E[t]$ and the variance σ_t as (Cohen, 1995)

$$E[t] = \int_{-\infty}^{+\infty} t\,|f(t)|^2 dt \tag{2.4}$$

$$\sigma_t^2 = \int_{-\infty}^{+\infty} \left(t - E[t]\right)^2 |f(t)|^2 dt = E\left[t^2\right] - E^2[t] \tag{2.5}$$

In a similar way, the average frequency $E[\omega]$ and the standard deviation σ_ω offer information about the frequency localization and bandwidth of the signal, i.e.:

$$E[\omega] = \int_{-\infty}^{+\infty} \omega \left|\hat{F}(\omega)\right|^2 d\omega \tag{2.6}$$

$$\sigma_\omega^2 = \int_{-\infty}^{+\infty} \left(\omega - E[\omega]\right)^2 \left|\hat{F}(\omega)\right|^2 d\omega = E\left[\omega^2\right] - E^2[\omega] \tag{2.7}$$

The use of sinusoidal functions for signal analysis achieves the best possible localization in frequency, since their frequency domain representation is a delta function. Consequently, the FT offers optimal resolution of the frequency content of a signal. However, the sinusoids have infinite support in time, thus they do not possess any time localization properties. Therefore, the FT of a signal does not provide information on the variation in time of frequency components (Newland, 1984; Cohen, 1995). For example, consider a signal obtained from the superposition of two harmonics with different amplitudes A_1 and A_2 and durations t_1 and t_2 plotted in the top panel of Figure 2.1:

$$f(t) = A_1 \sin \omega_1 t_1 + A_2 \sin \omega_2 t_2$$

$$\begin{array}{lll} A_1 = 0.7 & \omega_1 = 25 & t_1 \in [0,1] \\ A_2 = 0.5 & \omega_2 = 125 & t_2 \in [0.3, 0.7] \end{array} \tag{2.8}$$

Note that the in the Fourier amplitude spectrum (Figure 2.1– bottom panel) shows the frequency components in the signal without any information on their location or duration in time. In case this information is important, the Short Time Fourier Transform can be employed for analysis, as detailed in the following section.

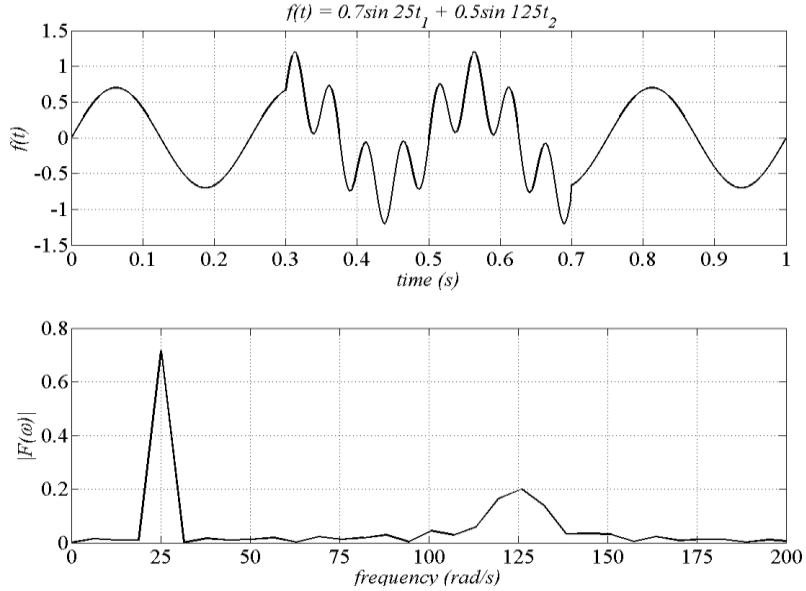

Figure 2.1. Superposition of two sinusoids with different durations and the corresponding FT magnitude spectrum

2.3. THE SHORT TIME FOURIER TRANSFORM

The Short Time Fourier Transform (STFT) computes the FT for successive parts of the signal, delimited by means of a window function *w(t)*. By sliding the window along the time axis using a translation parameter *b*, a representation of the frequency content of the signal localized in time is obtained as

$$STFT(t,\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(b) w(t-b) e^{-i\omega b} db \tag{2.9}$$

The localization in time is ensured through the following property of the window function *w(t)*:

$$w(t-b)=\begin{cases}1 & for\ t\ around\ b\\ 0 & otherwise\end{cases} \tag{2.10}$$

The distribution of the signal energy on the time-frequency plane can be approximated using the STFT by considering the so-called spectrogram, defined as:

$$S(t,\omega)=\left|STFT(t,\omega)\right|^2=\left|\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(b)w(t-b)e^{-i\omega b}db\right|^2 \tag{2.11}$$

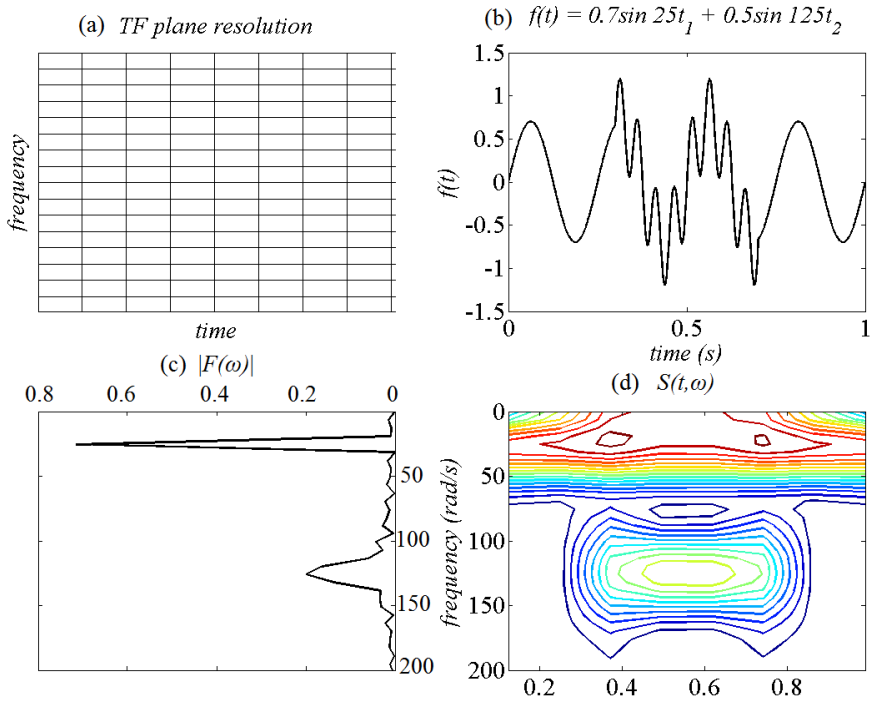

Figure 2.2. Spectrogram of a signal: (a) time-frequency plane discretization; (b) time-domain representation of the signal; (c) Fourier amplitude spectrum; (d) spectrogram of the signal

In Figure 2.2 the spectrogram of the signal defined in Eq. (2.8) is illustrated. Both frequency components and their duration in time can be identified in the spectrogram. However some differences appear in comparison with the FT magnitude spectrum: although different in magnitude, the frequency components appear to have similar

bandwidth. This dissimilarity is the effect of the resolution (i.e. the level of detail) employed for the representation. The resolution depends on the localization properties of the analysing window *w(t),* namely the duration (which can be measured using Eq. (2.5)) and the bandwidth of the windowed signal (Eq. (2.7)). Ideally, the smaller σ_t and σ_ω, the more accurate is the representation. However, the sizes of these quantities are interdependent and governed by the following equation:

$$\sigma_t \cdot \sigma_\omega \geq \frac{1}{2} \tag{2.12}$$

This dependency, characteristic to all oscillatory functions, implies a trade-off between the resolution in time and in frequency domain and is due to the Heisenberg (or uncertainty) principle (Cohen, 1995). The limitations posed by this principle are very intuitive: if a function is contracted in time (i.e. implying that it has good time localization), its frequency content broadens (i.e. reduced frequency localization); if it is dilated, the frequency content becomes better localized.

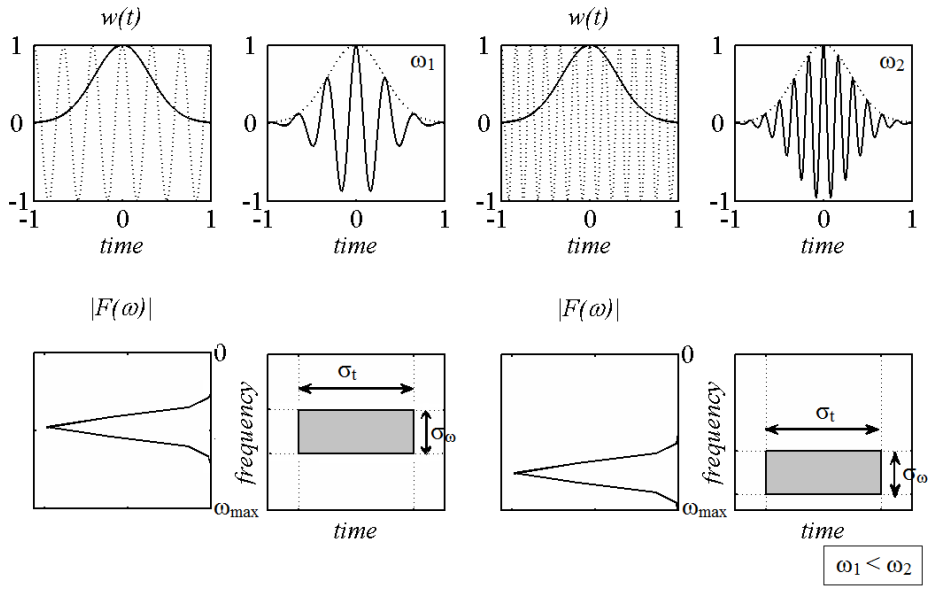

Figure 2.3.The effect of the fixed width window function on the accuracy of the time-frequency representation

In the case of the STFT the size of the window is chosen before the analysis and remains fixed at all times and frequencies; the result is a constant resolution across the time-frequency plane (Figure 2.2a). In Figure 2.3 the effect of a constant resolution across the time-frequency representation is exemplified. Note that as the frequency content analysed increases (thus the duration of the oscillations shortens) more oscillations are located in the same time-interval σ_t, leading to a less accurate localization of their occurrence in time.

The selection of a window-size capable to ensure satisfactory representation of the signal on the time-frequency plane is a challenging issue. This becomes even more critical when dealing with signals characterized by broad frequency content, as it is the case of those encountered in earthquake engineering. The wavelet transform emerged as a technique capable to overcome these limitations by using windows of varying sizes, which adapt to the frequency content being analysed.

2.4. THE WAVELET TRANSFORM

2.4.1. General remarks

The wavelet transform (WT) is a decomposing technique which employs short duration oscillatory functions, named wavelets, for projecting signals on the time-frequency plane (Daubechies, 1992; Mallat, 1989, 2009). A collection of wavelets is generally obtained using a given oscillatory function known as mother wavelet $\psi(t)$. The "daughter wavelets" are generated by scaling the mother wavelet, i.e. dilating or contracting its support in time, using a scaling factor $a > 0$ and without changing the number of oscillations, as shown in Figure 2.4. The daughter wavelets $\psi_{b,a}(t)$, normalized by a factor I_{ψ}^{a}, are translated along the time-axis of the signal by means of a parameter b:

$$\psi_{b,a}(t) = \frac{1}{I_\psi^a}\psi\left(\frac{t-b}{a}\right);\ a>0, b\in\mathbb{R} \tag{2.13}$$

Having a family of wavelets which satisfy specific conditions (detailed in the following section), the representation of the signal on the joint time-frequency plane is obtained by successively performing the inner product with each wavelet in the following way:

$$WT_\psi(b,a) = \frac{1}{I_\psi^a}\int_{-\infty}^{+\infty} f(t)\psi_{b,a}^*(t)dt \tag{2.14}$$

The wavelet transform of the previously considered signal (Eq. (2.8)) can be seen in Figure 2.5. According to the convolution theorem, the following correspondence between the time and frequency domains exists for finite energy functions (Newland, 1984; Cohen, 1995):

$$(f*g)(t) \leftrightarrow \hat{F}(\omega)\hat{G}(\omega) \tag{2.15}$$

where the symbol '*' stands for convolution. Consequently, it is less computational demanding to perform the wavelet transform in the frequency domain. Additionally, for scaling the wavelet function, the following relationship exists between the time and the frequency domain (Figure 2.4):

$$f\left(\frac{t}{a}\right) \leftrightarrow |a|\hat{F}(a\omega), \quad a>0 \tag{2.16}$$

The wavelet coefficients are a function of the wavelet's *scale*, thus the representation of the signal is obtained on the time-scale plane. However, the representation on the time-frequency plane may be desirable, since it is related the notion of frequencies obtained by standard Fourier analysis. Such a representation can be obtained by employing the

following relationship between the dominant frequency of the mother wavelet ω_c (i.e. the peak of its FT representation) and the frequency of the scaled wavelets (Teolis, 1998; Qian, 2002):

$$\omega_{\psi_{b,a}} = \frac{\omega_c}{a} \tag{2.17}$$

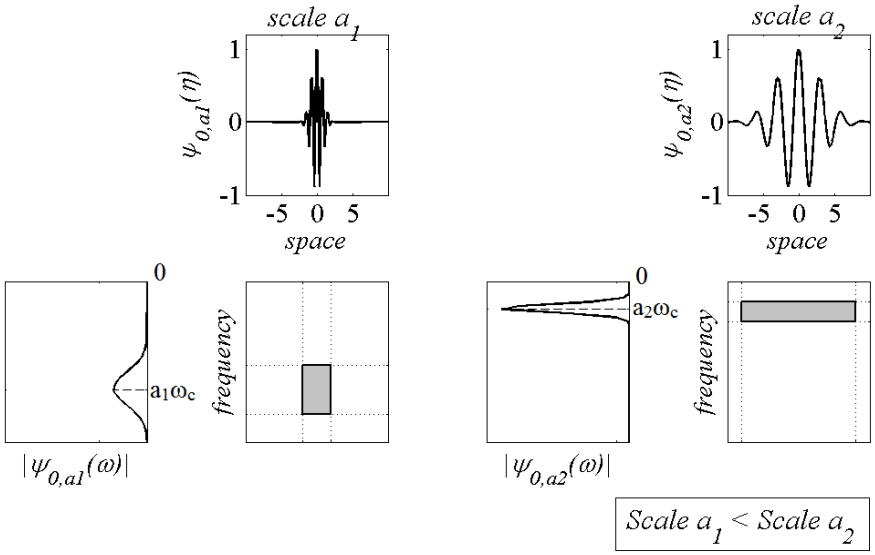

Figure 2.4. Time-domain and frequency domain representation of wavelet functions obtained for different scaling factors

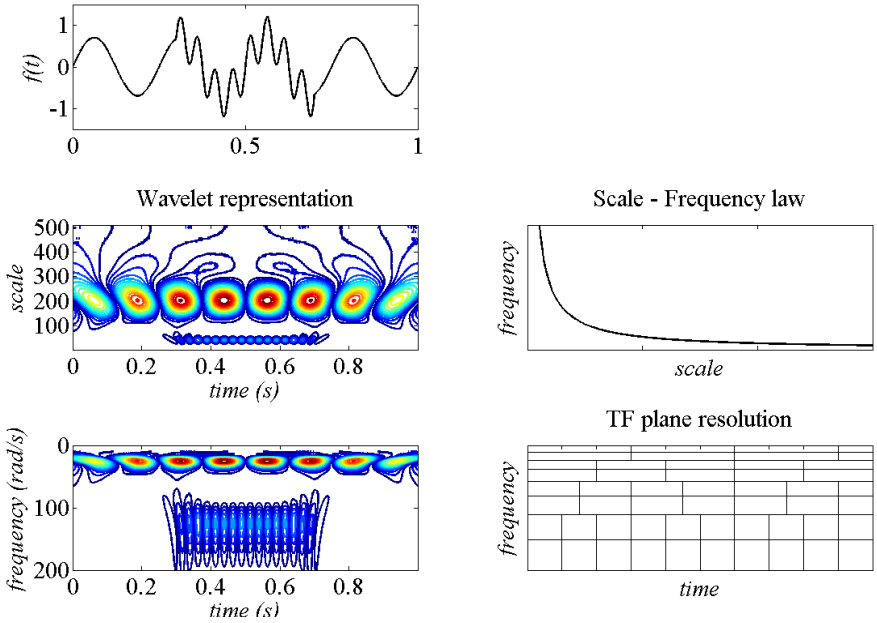

Figure 2.5. Wavelet transform: time-scale and corresponding time-frequency representations of a signal

2.4.2. Properties for wavelet functions

In order to ensure the reconstruction of the signal from the decomposition, wavelet functions must satisfy the following admissibility condition (Farge, 1992; Teolis, 1998; Qian, 2002; Mallat, 2009):

$$C_{\psi} = \int_{-\infty}^{+\infty} \frac{\left|\hat{\Psi}(\omega)\right|^2}{|\omega|} d\omega < \infty \tag{2.18}$$

where $\Psi(\omega)$ is the frequency domain representation of the mother wavelet. Equation (2.18) implies the function has finite energy; for this to be true, the wavelet has to have finite duration in time:

$$\int_{-\infty}^{+\infty} \left|\psi(t)\right| dt < \infty \tag{2.19}$$

A second implication is that the following equation needs to be satisfied as well:

$$\int_{0}^{+\infty} \frac{\left|\hat{\Psi}(\omega)\right|}{|\omega|} d\omega < \infty \tag{2.20}$$

The integral (2.20) is finite when the function $\left|\hat{\Psi}(\omega)\right|$ decays fast as $\omega \to 0$, which indicates that $\psi(t)$ needs to have zero-mean and thus it has to be an oscillatory function:

$$\langle \psi(0) \rangle = \int_{-\infty}^{+\infty} \psi(t) dt = 0 \tag{2.21}$$

2.4.3. Wavelet normalization and reconstruction of the decomposed signals

The magnitudes of the WT coefficients obtained at different scales have to be directly comparable, thus appropriate normalization of the wavelets needs to be performed. It is a common practice for the normalization factor I_{ψ}^{a} to ensure that wavelets at different scales have unit energy content (L_2 normalization) i.e.

$$\int_{-\infty}^{+\infty}\left|\frac{\psi_{b,a}(t)}{I_{\psi}^{a}}\right|^{2}dt=\int_{-\infty}^{+\infty}\left|\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)\right|^{2}dt=\left\|\psi_{b,a}(t)\right\|_{2}=1,\ (\forall)a>0 \tag{2.22}$$

This type of normalization (also implemented in the MATLAB Wavelet Toolbox) is appealing since it guarantees the conservation of the signal's total energy throughout the transform (Farge, 1992). In this context, the time-scale representation of the signal's energy content distribution, known as the scalogram, can be obtained as

$$Scal\left(b,a\right)=\left|WT\left(b,a\right)\right|^{2}=\left|\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}f\left(t\right)\psi_{b,a}^{*}(t)dt\right|^{2} \tag{2.23}$$

Further, the signal can be reconstructed from the transform by means of the following equation:

$$f\left(t\right)=C_{\psi}\iint\frac{1}{a^{2}}WT_{\psi}\left(a,b\right)\psi\left(\frac{t-b}{a}\right)dadb \tag{2.24}$$

Alternatively, different normalizations of the wavelet function can be performed. The wavelets can be normalized to have unit area (L_1 normalization):

$$\int_{-\infty}^{+\infty}\left|\frac{\psi_{b,a}(t)}{I_{\psi}^{a}}\right|dt=\int_{-\infty}^{+\infty}\left|\frac{1}{a}\psi\left(\frac{t-b}{a}\right)\right|dt=\left\|\psi_{b,a}(t)\right\|_{1}=1,\ \ (\forall)a>0 \tag{2.25}$$

or unit amplitude (L_∞ norm):

$$\left\|\psi_{b,a}(t)\right\|_\infty = 1, (\forall) a > 0 \tag{2.26}$$

The type of normalization used for the wavelet function influences the appearance of the time-frequency image of the signal, emphasizing differently the component frequency content, as observed and discussed in the comparative studies such as Farge (1992), Ventosa et al. (2008) or Vassiliou & Makris (2011).

2.4.4. Choice of the analysing wavelet function

The characteristics of the wavelet functions also influence the image of the signal obtained through the decomposition and should be considered when interpreting the resulting representations (Farge, 1992; Torrence & Compo, 1998):

- Complex versus real wavelets

Complex wavelets offer information about the phase and the amplitude which makes them suitable for the analysis of highly oscillatory functions. Real wavelets return information about the amplitude and are useful for identifying peaks or discontinuities in the signals.

- Orthogonal versus non-orthogonal wavelets

Orthogonal wavelets are families of uncorrelated functions, with the zero inner product between any two different functions:

$$\left\langle \psi_{b,a_1}, \psi_{b,a_2} \right\rangle = \int \psi^*_{b,a_1}(t)\psi_{b,a_2}(t)\,dt = 0, \qquad for\ a_1 \neq a_2; a_1, a_2 > 0 \tag{2.27}$$

The orthogonal WT offers the most compact representation of a signal. The number of convolutions is proportional to the scale of the wavelet and the WT coefficients are

independent. The non-orthogonal wavelets lead to over-complete representations, which are suitable for smooth representations of signals.

- Width of the wavelet

The width is relevant in terms of localization of the frequency content in time and frequency. Each wavelet defines a window on the time-frequency plane with the following area (Qian, 2002):

$$\left[t-a\sigma_t, t+a\sigma_t\right]\times\left[\omega-\frac{\sigma_\omega}{a}, \omega+\frac{\sigma_\omega}{a}\right] \tag{2.28}$$

$$2a\sigma_t \cdot 2\frac{\sigma_\omega}{a} = 4\sigma_t\sigma_\omega = const \tag{2.29}$$

Although both WT and STFT are subjected to the limitations of the uncertainty principle, the difference between them lays in the fact that the resolution (i.e. level of detail) of the WT representation can be varied according to the range of frequencies analysed (as it can be seen by comparing Figure 2.2 and Figure 2.5), as long as the area of each box delimited by the wavelet on the time-frequency plane remains constant (Figure 2.4).

- Shape of the wavelet

This attribute refers to the smoothness of the function. The choice of the shape depends on the visual appearance of the signal – for smooth signals smooth functions are preferred, while for signals with sharp jumps, other functions might be more appropriate (Daubechies, 1992).

From the numerous functions existing in the literature, the following mother wavelets are often encountered in earthquake-related applications: the Daubechies wavelets (Baker, 2007; Yaghmaei-Sabegh, 2010; Mollaioli & Bosi, 2012), the Coiflets (Todorovska et al., 2009; Yaghmaei-Sabegh, 2010), second or third order derivatives of

the Gaussian (Vassiliou & Makris, 2011; Gupta & Mukhopadhyay, 2013), Meyer wavelets (Yamamoto & Baker, 2013), harmonic wavelets (Spanos & Failla, 2004; Giaralis & Spanos, 2009; Spanos et al., 2009; Spanos & Kougioumtzoglou, 2012).

2.4.5. Types of wavelet transform

There are several ways to perform the WT depending on the scope of the analysis (Daubechies, 1992). The difference between them lies in the values adopted for the scaling and translation parameters and in the characteristics of the mother wavelet. The types of WT used in this work are briefly reviewed herein, while a more comprehensive treatment of this topic can be found in (Daubechies, 1992; Teolis, 1998; Mallat, 2009).

The *continuous wavelet transform* (CWT) is commonly used for achieving smooth representations of the signals on the time-frequency plane. The dilation and translation parameters a and b vary continuously over the time-frequency plane, leading to highly redundant representations, and thus very detailed portrayals of the signal's content evolution over the time-frequency plane (see Figure 2.5). The CWT is a computationally demanding technique; its numerical implementation involves considering specific discrete values, whose density determines the resolution of the output.

The *discrete wavelet transform* (DWT) is a complete and non-redundant version of the WT usually employed for data compression, for modelling purposes or for denoising (Teolis, 1998; Mallat, 2009). The wavelet families used for performing DWT form bases of orthogonal functions. The values of the scaling and translation parameters commonly follow a dyadic sampling, i.e.:

$$\begin{aligned} a &= \frac{1}{2^i},\ i \in \mathbb{Z} - \{0\} \\ b &= k2^i,\ k \in \mathbb{Z} \end{aligned} \tag{2.30}$$

From a numerical implementation point of view, an alternative and more straightforward approach than successively applying Equation (2.14) can be used for obtaining the coefficients of the DWT, namely by using filter banks (see Daubechies, 1992; Mallat, 1989, 2009). Mallat (1989, 2009) showed that the decomposition of a signal on a basis of wavelets consisting of compactly supported functions sampled on a dyadic grid, is similar to a repeated filtering of the signals using conjugate mirror filters (used in filter banks). Based on this property, appropriately chosen filters can be used to decompose a signal on the TF plane, following the methodology detailed further on.

Filter banks are used in signal processing for obtaining different level of approximations of the analysed input. The signal is passed through a pair of filters, a low-pass (LP) and high-pass (HP), which separate its content in approximations and details, as it can be seen in the left panel of Figure 2.6. The approximation coefficients (A_{ji}) are the output of the LP filter and represent averages of the signal. The detail coefficients (D_{ji}) correspond to the HP filter and represent its oscillatory parts, i.e. the details which are lost through averaging (Mallat, 1989, 2009).

The filtering is performed successively to the output of the LP filter. At each step, the down-sampling ($\downarrow 2$) of the filtered signal ensures the non-redundancy of the results. The procedure is repeated until the desired level of approximation is obtained. For each level j of the decomposition, the output of the LP filter is $A_{j+1} = A_j + D_j$. This implies that the original signal can be reconstructed by adding up the final output of the LP filter with all the detail coefficients obtained, i.e.:

$$f(t) = A_n + \sum_{j=1}^{n} D_j \tag{2.31}$$

In the case of the DWT, this successive filtering results in a wavelet tree (Figure 2.6). Notice that since only the approximations of the signal are further filtered, the wavelet

tree is asymmetric. The number of times the filters are applied gives the depth of the tree. Since in practice we work with discrete-valued signals, the maximum depth of the tree (J) is limited by the total number of samples (discrete points) N in a signal (Mallat, 2009):

$$J = \log_2 N \tag{2.32}$$

There are cases when a more uniform resolution of the TF plane is needed. This can be achieved by applying the *wavelet packets transform* (WPT), an extended version of the DWT. In this case, not only the approximations of the signal, but also the detail coefficients are further processed until the sought level of approximation is acquired, as shown in the right panel of Figure 2.6. The wavelet tree obtained in this case is symmetrical and just like in the previous case the nodes are orthogonal to each other, encompassing information from adjacent frequency bands. The decomposition is still non-redundant, since the output of each filter and at each level is down-sampled before being further processed. For cases when a more detailed discretization is desired in specific frequency intervals, the output of the corresponding nodes is further processed. From a TF representation perspective, the flexibility of the WPT is important since it offers the opportunity to "zoom-in" within certain frequency bands at will.

In the following sections the families of wavelets considered for the purpose of this work are presented: the generalized harmonic wavelets and the Meyer wavelet packets.

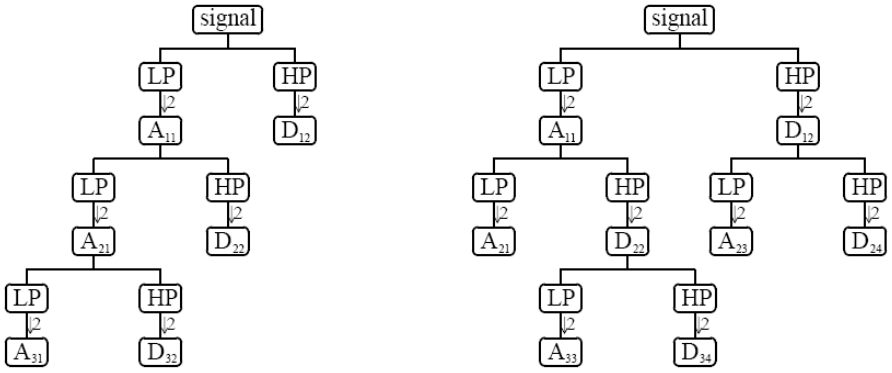

Figure 2.6. Wavelet tree for DWT(left) and for WPT (right)

2.4.6. Generalized harmonic wavelet transform

The generalized harmonic wavelets (GHW), introduced by Newland (1994), are complex functions characterized by a box-like shape in the frequency domain. Their frequency support is defined by means of two parameters: the lower limit *m* and the upper limit *n* (see Figure 2.7). The Fourier domain representation of GHW at scale *(m, n)* and located at the time instant *b* is given by the formula:

$$\hat{\psi}_{m,n}^{b}(\omega)=\begin{cases}\dfrac{1}{2\pi(n-m)}, & (2\pi)m\le\omega\le(2\pi)n, \qquad n>m\\ 0, & otherwise\end{cases} \tag{2.33}$$

The rate of decay in time for the GHW is relatively slow, i.e. proportional with $1/t$, implying they have an average time-localization (Newland, 1994). The wavelets covering different frequency bands are orthogonal to each other. When the family of wavelets used for analysis is defined such that the functions are adjacent to each other and cover the entire frequency span (Figure 2.8) a complete and non-redundant representation of the signal is obtained.

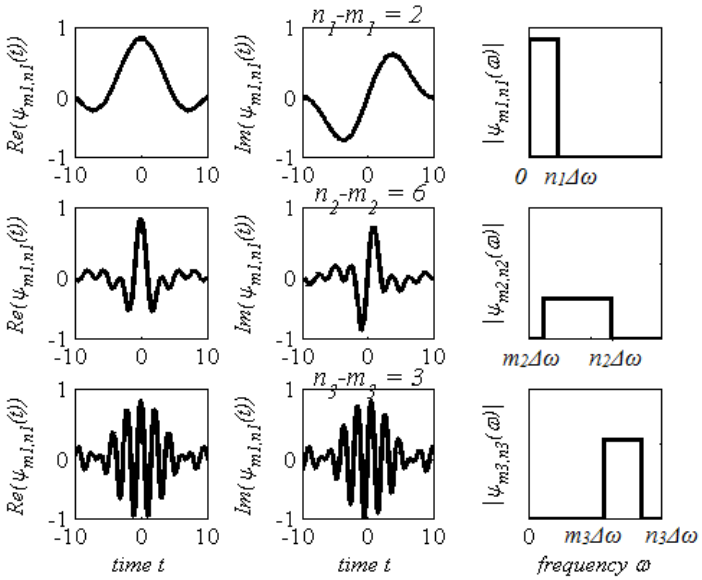

Figure 2.7. Generalized harmonic wavelets at different scales

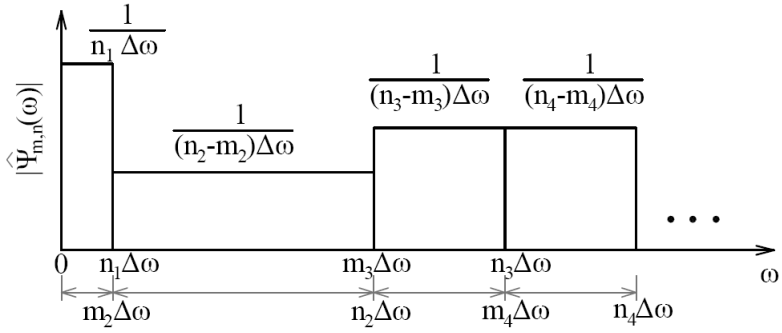

Figure 2.8. Generalized harmonic wavelet basis

The mathematical implementation of the generalized HWT is very straight forward, thanks to their characteristic shape. The Fourier coefficients of the signal are separated in blocks corresponding to each wavelet in the family and by taking the inverse FFT of each frequency band, the GHW transform coefficients are obtained (see Figure 2.9, left panel). In this sense the GHW transform is similar to the wavelet packets decomposition since it filters parts of the signal with complementary frequency content; however, due to the way the wavelet functions are defined it is not bounded to a dyadic discretization and this allows for more freedom in selecting the discretization for the time-frequency plane.

The GHWs are used in the following chapters in a CWT context, in order to obtain a complete and detailed characterization of the signals (Newland, 1999). To obtain the CWT, the Fourier product between each wavelet and the signal is padded with zeros until the entire frequency axis is covered, before taking the IFFT to obtain the wavelet transform coefficients (Figure 2.9 - right panel). The zero-padding of the frequency domain product between the signal and the wavelet is equivalent with an interpolation between the coefficients for the intermediate time instants, leading to a smoother and more detailed image of the signal's content (Newland, 1999).

Due to the flexibility in adjusting the resolution of the TF representation, the GHWs have been used in seismic related applications, i.e. for power spectrum estimation (Spanos & Failla, 2004; Spanos & Kougioumtzoglou, 2012) or for response spectrum matching modification (Giaralis & Spanos, 2009; Spanos et al., 2009).

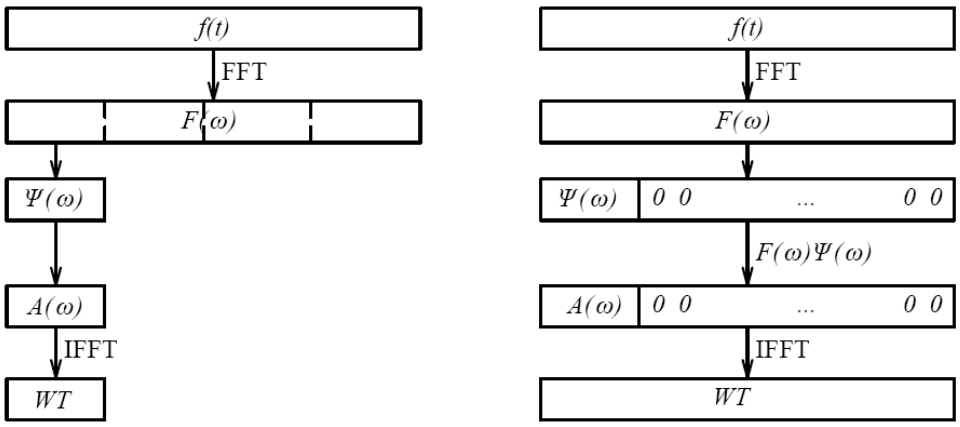

Figure 2.9. Discrete GHWT (left) and continuous GHWT (right)

2.4.7. Meyer wavelet packets transform

The Meyer wavelets (MW), plotted in Figure 2.10, are orthogonal functions with smooth shape in the frequency domain, which ensures a faster decay in time in comparison with the harmonic wavelets. They are constructed using conjugate mirror

filters and form bases, although overlapping exists between adjacent scales (Daubechies, 1992; Mallat, 2009). The Fourier transform of MW is defined as

$$\hat{\psi}(\omega)=\begin{cases}\frac{1}{\sqrt{2\pi}}e^{\frac{i\omega}{2}}\sin\left[\frac{\pi}{2}\nu\left(\frac{3}{2\pi}|\omega|-1\right)\right], \frac{2\pi}{3}\le|\omega|\le\frac{4\pi}{3}\\ \frac{1}{\sqrt{2\pi}}e^{\frac{i\omega}{2}}\cos\left[\frac{\pi}{2}\nu\left(\frac{3}{4\pi}|\omega|-1\right)\right], \frac{4\pi}{3}\le|\omega|\le\frac{4\pi}{3}\\ 0, otherwise\end{cases}, \tag{2.34}$$

where the function ν satisfies the conditions

$$\nu(u)=\begin{cases}0, u\le 0\\ 1, u\ge 0\end{cases} \quad and \quad \nu(u)+\nu(1-u)=1. \tag{2.35}$$

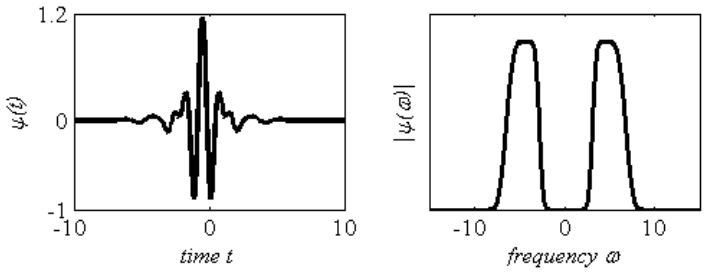

Figure 2.10. Meyer wavelet: time domain and frequency domain representation

In this work the MW are used in a WPT context, following the algorithm presented in Section 2.4.5. By plotting the coefficients of the transform a TF image of the signal's content with a dyadic resolution is obtained. The MW packets have been previously employed in earthquake engineering applications due to their localization properties and orthogonality (i.e. Yamamoto & Baker 2011, 2013).

2.5. THE S-TRANSFORM

The S-transform (ST) was developed by Stockwell et al. (1996) as a combination between the STFT and the CWT. Recalling equation (2.9), consider the following Gaussian window function, normalized to unit area, whose width and location can be varied by means of the parameters *a* and *b*

$$w(b,a)=\frac{1}{a\sqrt{2\pi}}e^{-\left(\frac{t-b}{2a}\right)^2} \tag{2.36}$$

The equation (2.9) becomes

$$ST(b,a,\omega_i)=\int_{-\infty}^{+\infty} f(t)\frac{1}{a\sqrt{2\pi}}e^{-\left(\frac{t-b}{2a}\right)^2}e^{-i\omega_i t}dt \tag{2.37}$$

The time-width of the Gaussian window can be defined as a function of the oscillatory part frequency, i.e. $a=2\pi/|\omega_i|$, which leads to the following expression for the ST:

$$ST(b,\omega_i)=\frac{|\omega_i|}{(2\pi)^{3/2}}\int_{-\infty}^{+\infty} f(t)e^{-\frac{\omega_i^2(t-b)^2}{2\cdot(2\pi)^2}}e^{-i\omega_i t}dt=\frac{1}{I_\psi^a}\int_{-\infty}^{+\infty} f(t)\psi_{b,\omega_i}(t)dt \tag{2.38}$$

$$where \quad \psi_{b,\omega_i}(t)=e^{-\frac{\omega_i^2(t-b)^2}{2\cdot(2\pi)^2}}e^{-i\omega_i t} \tag{2.39}$$

The analysing function $\psi_{b,\omega_i}(t)$ in Eq. (2.39) resembles to the Morlet wavelet family normalized to L_1 norm (i.e. unit area), which is given by the following equation:

$$\psi_{b,a}(t)=\frac{1}{a\sqrt{2\pi}}e^{-\frac{(t-b)^2}{2a^2}}e^{-i\omega_0\left(\frac{t-b}{a}\right)} \tag{2.40}$$

This type of normalization ensures a direct relationship with the Fourier amplitude, but does not ensure the energy conservation (Ventosa et al, 2008), as detailed in Section 2.4.3.

By analogy there is a clear similarity between equations (2.14) and (2.38); however, there are certain differences to note. First, the CWT uses the notion of scale which has an indirect relation to the frequency depending on the wavelet frequency – scale law, while the ST is defined as a function of frequency. Secondly, when generating the Morlet wavelet family the phase of the oscillatory part is continuously varied by means of the translation parameter *b*. In the case of the ST, the phase of the harmonic remains fixed as it can be observed in Figure 2.11. As a result, the localizations of the amplitude and phase spectra are independent from each other, property known as absolutely referenced (i.e. "fixed") phase information (Stockwell et al., 1996).

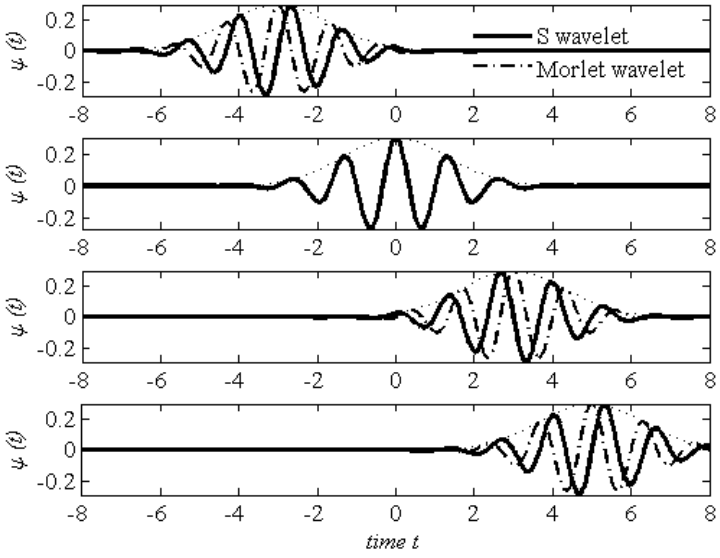

Figure 2.11. The difference between the S-wavelet and the Morlet wavelet: in the case of S-Transform the phase is absolutely referenced to the initial point.

Concerning the analysing wavelets for the ST, there are no restrictions on the mean value for these functions. While the zero-mean condition must be obeyed in the case of CWT in order to ensure the original signal is recoverable, for the ST the reconstruction

is secured through to the direct relationship with the Fourier spectrum (Ventosa et al, 2008):

$$f(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} ST_f(b,\omega)dbe^{i2\pi ft}d\omega \tag{2.41}$$

Several windows have been proposed for the ST in order to improve its adaptability to various applications. McFadden et al. (1999) modified the Gaussian window by adding an exponential tail after the peak is attained. According to the authors, this generalized ST can accommodate any suitable function, without any restrictions on the symmetry. Stockwell (1999) generalized the S-transform by ensuring more flexibility in the choice of the dilation parameter by means of the parameter k, which can take various values i.e.

$$a = \frac{k}{f} \tag{2.42}$$

Pinnegar & Mansinha (2004) introduced a symmetrical window which incorporates an extra parameter controlling the phase modulation and connecting the S-transform with the chirplet transform.

The S-transform has been applied to a wide range of applications, from detecting damage (faults) in machineries (McFadden et al., 1999) to denoising seismic signals (Pinnegar & Eaton, 2003; Askari & Siahkoohi, 2008; Parolai, 2009).

2.6. THE EMPIRICAL MODE DECOMPOSITION

The empirical mode decomposition (EMD) is a non-parametric, data-driven technique introduced by Huang et al. (1998) which allows for the breakdown of signals in fundamental components, known as intrinsic mode functions (IMFs). The IMFs are generally defined as functions which satisfy the following conditions: (i) at each time

instant they have zero-mean and (ii) the number of extremes is equal to / or differs by one from the number of zero-crossings. An example of the EMD performed for a simple signal is presented in Figure 2.12.

The scope of the EMD is to decompose highly non-stationary signals into simpler functions, which are characterized by a unique frequency component at each time instant (i.e. mono-component signals). This leads to a clear and straight forward characterization of the underlying frequency content distribution and evolution, which can be obtained by applying the Hilbert transform to the resulting IMFs (Huang et al, 1998).

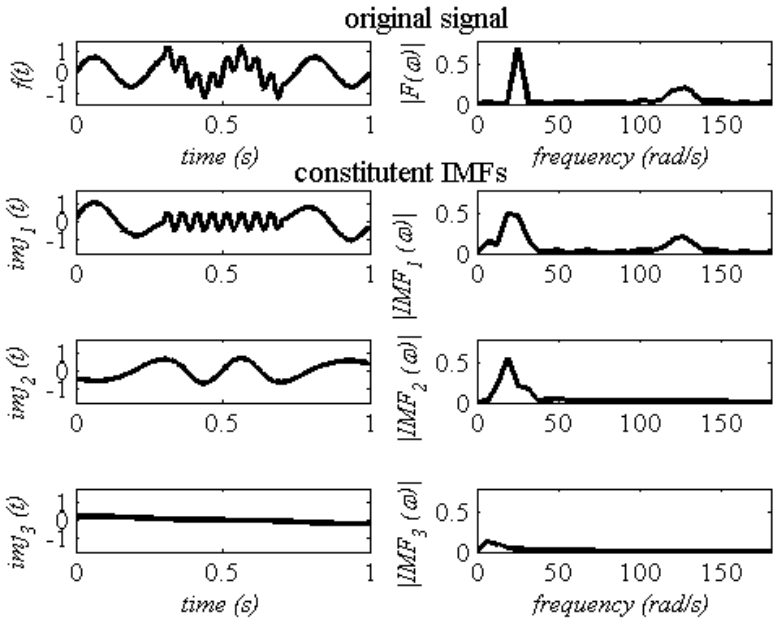

Figure 2.12. Intrinsic mode functions (IMFs) and their Fourier transform coefficients of a signal

The decomposition of a signal *f(t)* into IMFs can be obtain through the following steps (summarized in Figure 2.13):

1. Set $h_0(t) = f(t)$.
2. Identify the j^{th} IMF:
    a. Identify the local extremes of the time-series;
    b. Compute the upper $e_{up}(t)$ and lower $e_{lo}(t)$ envelopes, usually by cubic spline interpolation;

c. Compute the mid-point between the envelopes and subtract it from the data:

$$m(t) = \frac{e_{up}(t) - e_{lo}(t)}{2}$$
$$h_j(t) = h_{j-1}(t) - m(t) \tag{2.43}$$

d. Check if the IMF criteria are met; if yes $IMF_i(t) = h_j(t)$; otherwise repeat steps a-c until the IMF criteria are met;

3. Obtain the residual signal by extracting the IMF from the data

$$r_i(t) = h_0(t) - IMF_i(t) \tag{2.44}$$

4. Repeat the procedure until the residual becomes a monotonic function or has maximum two extreme values.

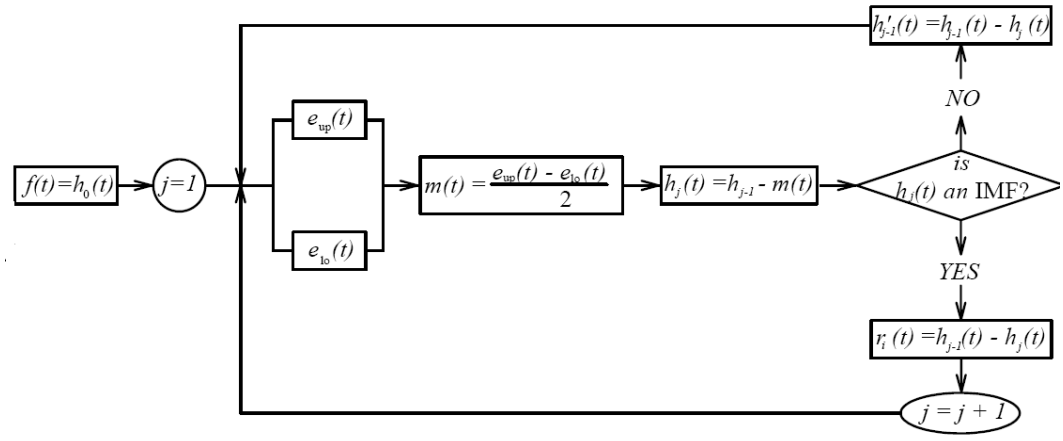

Figure 2.13. Empirical mode decomposition scheme

The number of IMFs obtained depends on the signal analysed, but it is also influenced by specific choices in the algorithm, such as the stopping criteria or the type of function used for enveloping (Huang et al, 2003; Dätig & Schlurmann, 2004). For highly non-stationary signals, as is the case of those encountered in earthquake engineering, the

steps 2.a.- 2.c. need to be repeated several times in order to obtain adequate IMFs. The role of this process is to avoid having two frequency components at the same instant and to obtain more symmetrical waveforms. This refining procedure is known as sifting and continues until pre-set stopping criteria like are met. Such stopping criteria can be thresholds for the amplitude of *h(t)* or a fixed number of siftings (Huang et al, 1998; Huang et al, 2003; Rilling et al., 2003). Caution should be exercised when choosing the stopping criteria: they have to be strict enough to ensure the components are realistically separated, but on the other hand they have to be flexible enough not to cause smoothing of the data and thus lead to unrealistic results (Huang et al, 2003).

The EMD decomposes the signal into a small number of intrinsic components compared to the FFT or the WT. The resulting IMFs are locally orthogonal to each other, however global orthogonality is not necessarily ensured, as it can be observed (Figure 2.12). From this perspective, the algorithm has been compared with an overlapping filter bank (Flandrin et al., 2004). Although mono-component, the resulting IMFs are not stationary, their frequency and amplitude varying over time. The high adaptability to the data and the intuitive, simple algorithm used for the EMD attracted the interest of researchers from different fields. On the other hand, due to the fact that it is based on an empirical algorithm rather than a mathematical formulation, any statistical processing of the results is quite challenging, limiting the use of the EMD mostly for characterization of specific records, i.e. seismic signals (i.e. Loh et al., 2001; Huang et al, 2001; Zhang et al., 2003; Spanos et al., 2007; Yinfeng et al., 2008), water waves (Dätig & Schlurmann, 2004), wind data (Huang et al., 1998), for denoising (Guo et al., 2012) or for damage detection (Xu et al., 2010).

CHAPTER 3 : PULSE-LIKE GROUND MOTIONS CHARACTERIZATION, EXTRACTION AND SIMULATION

3.1. PRELIMNARY REMARKS

Pulse-like seismic ground motions are characterized by the presence of high amplitude, long period pulses, which severely influence the behaviour of a wide range of relatively flexible structures in the affected areas. Historically, pulse-like ground motion (PLGM) related research started with the 1952 Kern County (California) event, followed by the 1966 Parkfield and the 1971 San Fernando (California) earthquakes. However, more extensive research efforts were devoted after the 1994 Northridge (California), 1999 Kobe (Japan) and 1999 Izmit (Turkey), 1999 Chi Chi (Taiwan) events, which caused significant structural damage especially in the case of buildings with medium to long natural periods (Sommerville 1997, 1998, 2000, 2002; Moustafa & Takewaki, 2010; Mavroeidis & Papageorgiou, 2003, Tang & Zhang, 2011). Aiming to improve the structural behaviour in the regions likely to be subjected to such ground motions, but also for seismic risk assessment studies, there is a lot of interest currently focussing on this topic. Several themes were identified in the PLGM related research, namely (i) the identification of the physical conditions which favour their occurrence and (ii) the identification of measurable distinguishing features, based on which they can be (iii) modelled and then (iv) simulated for the purpose of seismic risk assessment.

A record-based comparison between a pulse-like and a pulse-free record is performed in the next section, in terms of energy distribution on the time-frequency plane, peak

values and structural impact in order to identify distinct features which are commonly associated with PLGMs. Further, the physical causes presented in the literature as sources for PLGMs are summarized and the parameters employed for pulse characterization are presented. The pulse models existing in the literature are also discussed, followed by the commonly used approaches for identification and extraction of pulses from records, which include:

(i) fitting simplified, deterministic pulse models to time-domain representation of the signal (Menun & Fu, 2002; Mavroeidis & Papageorgiou, 2003; Moustafa & Takewaki, 2010; Yaghmaei-Sabegh, 2010) or to their response spectra (Tang & Zhang, 2011);

(ii) applying time-frequency signal processing techniques to the acceleration or velocity traces (Zhang et al, 2005; Baker, 2007; Xu & Agrawal, 2010; Vassiliou & Makris, 2011);

(iii) filtering the low-frequency content of the signal (Ghahari et al., 2010; Mukhopadhyay & Gupta, 2013a).

In the final section of this chapter the modelling approaches employed for PLGMs representation and/or simulation are presented.

3.2. PHENOMENOLOGICAL AND PHYSICAL CONSIDERATIONS

Consider the time-history traces of two accelerograms recorded during similar magnitude earthquakes, at Imperial Valley, Southern California (Figure 3.1.a). One belongs to the 1940 El Centro event of moment magnitude 6.9, while the second was recorded during the 1979 Imperial Valley strong earthquake of magnitude 6.4 and was classified as being of pulse-type. They were both recorded in the proximity of a seismic fault and have comparable peak ground accelerations.

In Figure 3.1.a, the distribution of the energy content in time and in frequency (TFD) is displayed for each accelerogram. The TFDs are obtained using the continuous wavelet transform and employing the generalized harmonic wavelets as decomposing functions, following the methodology presented in Section 2.4.6. It can be observed that both are characterized by broadband frequency content. In the case of the pulse-free accelerogram belonging to the El Centro event, the distribution of the energy in frequency is relatively uniform over the interval ~[5, 15]rad/s, while for the pulse-like Imperial Valley record most of the energy is concentrated in the low-frequency interval under ~5 rad/s. A first distinguishing aspect for PLGMs is thus the concentration of a significant amount of the total energy released during the earthquake in a narrow and very low-frequency band.

Further, the plots of the velocity and displacement traces of the two records (Figure 3.1.b) show that although the pulse may not be readily distinguishable in the accelerograms, it can be easily identified in the velocity and displacement records. As a consequence of the pulse, the pulse-like record is characterized by higher ratios between the peak ground velocity/peak ground acceleration and peak ground displacement/peak ground acceleration in comparison with the pulse-free earthquake of similar intensity.

From a structural engineering perspective, these features are of interest because they result into significantly higher demands on a wide range of structures, as it is indicated by the corresponding response spectra (Figure 3.1.c). Malhotra (1999) argues that this is the result of a widening of the acceleration sensitive zone, which leads to a reduced apparent flexibility of buildings and consequently to the predominance of the first vibration mode in structural responses for a wider range of natural periods. It implies also that the efficiency of supplemental damping is significantly reduced. Latter studies showed increased base-shear demands and larger interstorey drifts in the structures

when subjected to pulse-like ground motions (e.g. Alavi & Krawinkler, 2001; Tothong & Cornell, 2008; Sehhati et al, 2011).

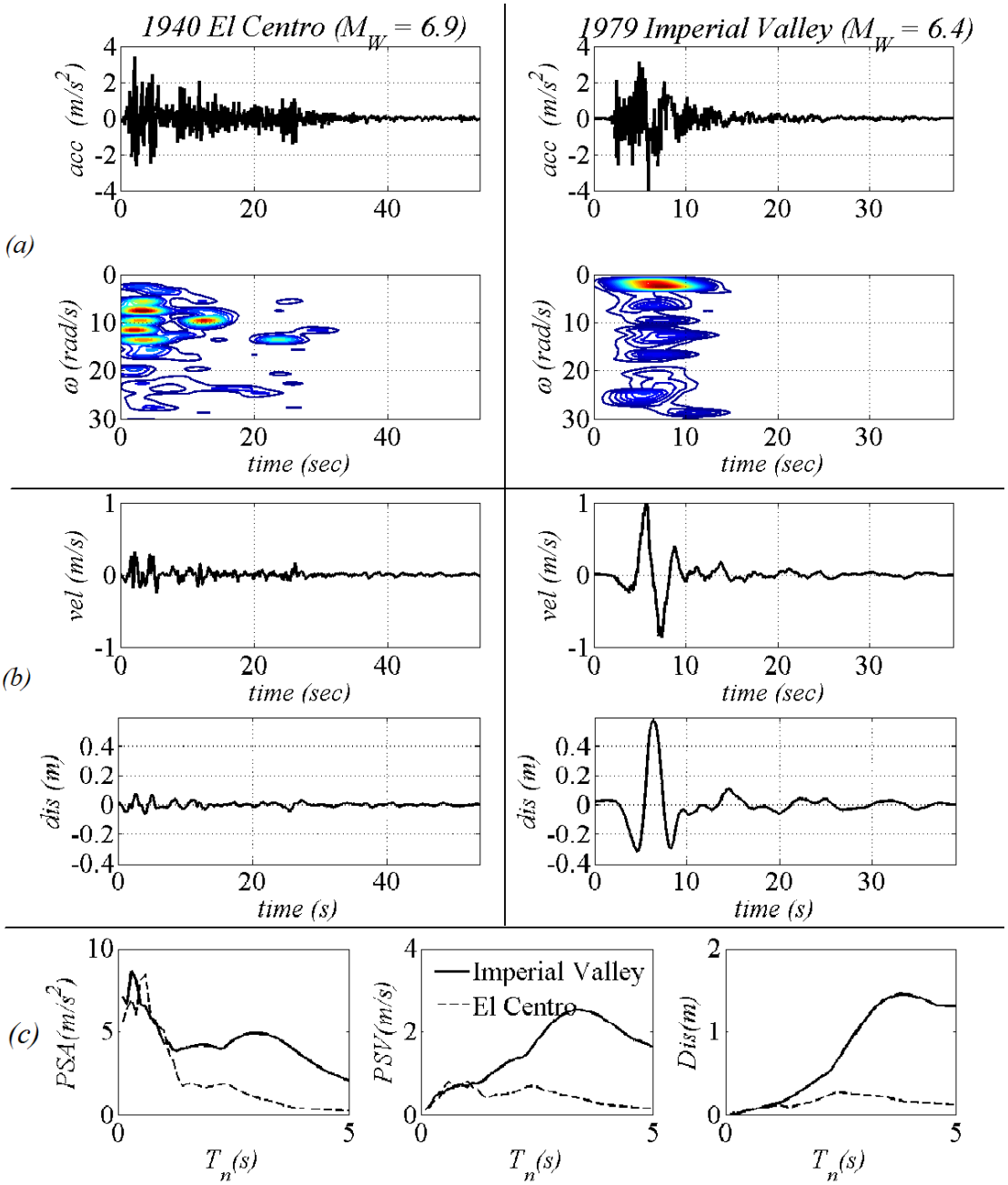

Figure 3.1. Non-pulse ground motion versus pulse-like ground motion of similar magnitudes and peak ground accelerations, recorded at Imperial Valley, Southern California: (a) accelerograms and their corresponding time-frequency energy distributions; (b) velocity and displacement time-histories; (c) 5% elastic response spectra

From a physical perspective, the presence of the strong low-frequency pulses in the ground motions is the result of a combination of various factors: the rupture mechanism, the source-to-site geometry, the displacement caused by the tectonic movement (temporary or permanent), asperities on the fault rupture or earth structure. However, it

has been observed that near-fault areas (i.e. located at 5-20(30) km from the seismic fault), are more likely to experience this type of ground motion as a consequence of the directivity effect or of the fling step effect. For this reason, pulse-like earthquakes are commonly referred to as near-field (fault) earthquakes, although it should be noted that not all the ground motions experienced in the proximity of seismic faults are of pulse-type (Iervolino & Cornell, 2008; Shahi & Baker, 2011).

The directivity effect is a dynamic phenomenon, caused by the tendency of the fault rupture to concentrate the wave energy along the fault. Depending on the position relatively to the fault, a specific site can experience "forward-directivity", when the rupture propagates towards the site or "backward-directivity", for ruptures propagating away from the site. The time-histories recorded as a consequence of forward-directivity (FD) effect have shorter duration and present large amplitude velocity pulses at the beginning of the records. The time-histories recorded in the case of backward-directivity have longer durations and small amplitudes (Sommerville, 1997; Dabaghi & Der Kiureghian, 2011).

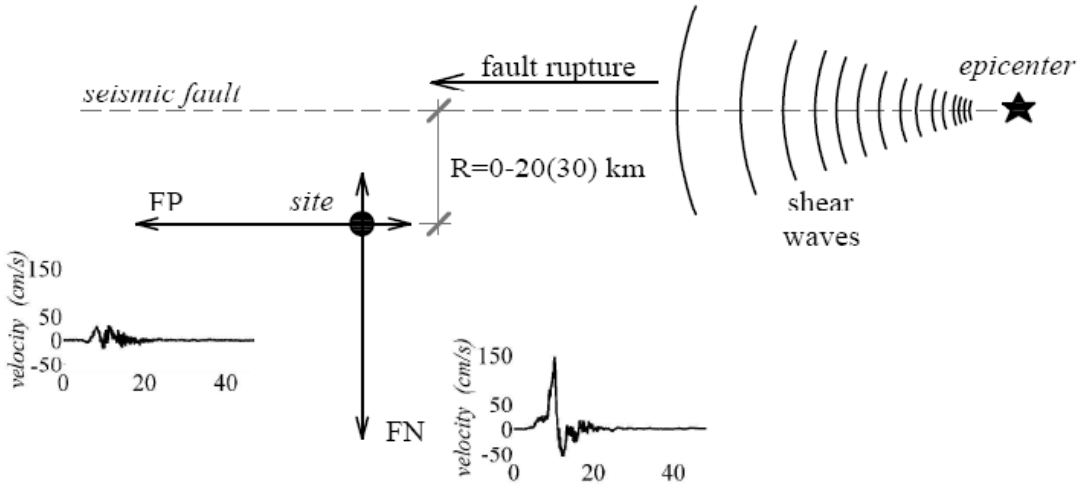

Figure 3.2. Schematization of the forward directivity effect (planar view)

The FD effect is the most common cause of strong pulses in seismic records and it attenuates as the distance from the fault increases. The pulses build-up when the rupture

and the shear waves propagate towards the site with similar velocity, causing all the seismic energy to arrive at once, as it is illustrated in Figure 3.2. If the direction of the waves travelling from the hypocentre to the site coincides with the direction of the slip on the fault, all the conditions necessary for the build-up of the FD effect are met. This happens usually in the case of strike-slip faults and, sometimes, dip-slip faults can experience this phenomenon also. The time-histories recorded for the fault-normal direction (FN) will present a number of two-sided oscillations and amplitudes significantly higher than those recorded in the fault parallel direction (FP), unlike in the case of far-field recorded time-histories where they have similar magnitude values. It is worth to be noted that not all the seismic ground-motions recorded in the proximity of the fault will exhibit FD effects; however the probability increases as the site is located closer to the fault (Sommerville, 2000; Iervolino & Cornell, 2008; Dabaghi & Der Kiureghian, 2011; NEHRP Consultants Joint Venture, 2011).

Another cause of pulses is the fling step (FS) effect, which causes permanent ground displacement and results in one-sided pulses in the velocity time-histories and non-zero final displacement values. The pulses caused by the FD effect are usually treated separately from those caused by the FS effect, due to their different features and source mechanisms (Bray & Rodriguez-Marek, 2004), and are not considered in this work.

3.3. CHARACTERIZATION OF PULSES

3.3.1. Parameters employed for pulse characterization

An important step in enhancing the understanding and evaluation of the structural impact of PLGMs consists in proposing approaches for quantifying their underlying properties. The commonly used parameters for this purpose are detailed herein, while some of them are illustrated in Figure 3.3.

The period of the pulse (T_p) is a fundamental parameter employed for characterization, since its relationship to the natural period of the structure gives information on the expected level of structural damage to be experienced during the ground motion. Its value depends on a combination of factors: the magnitude of the earthquake, the earth structure (rock or soil), the type of rupture (forward-directivity pulses are usually of higher amplitude and shorter period that those due to the fling-step effect), but also on the approach adopted for measuring the period (i.e. Mavroeidis & Papageorgiou, 2003, Bray & Rodriguez-Marek, 2004, Dabaghi et al., 2011). Currently there is no generally accepted way to measure the period and various alternatives are being used: the distance between two successive zero-crossings (Bray & Rodriguez-Marek, 2004), the period of a truncated sine-wave that fits the pulse (Menun & Fu, 2002), the time interval until the velocity decays at 10% from the peak value, the period associated with the peak value of the velocity pseudo-response spectrum (Alavi & Krawinkler, 2001; Mavroeidis & Papageorgiou, 2003) or as the dominant period of the wavelet corresponding to the largest coefficient of the wavelet transform, when applied to the velocity trace (Baker, 2007).

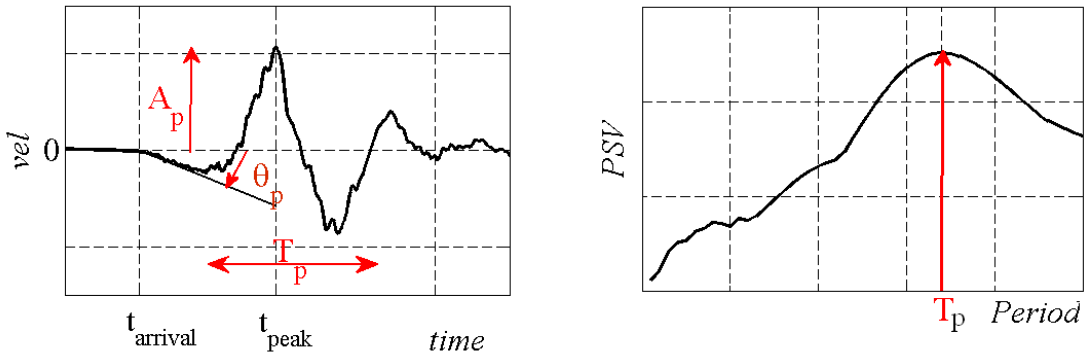

Figure 3.3. Parameters commonly employed for pulse characterization

Another main characteristic is the amplitude of the velocity pulse A_p, which depends on the magnitude, distance from the fault, soil and number of analysed records (e.g. Mavroeidis & Papageorgiou, 2003; Bray & Rodriguez-Marek, 2004).

The number of pulses varies with the slip distribution on the causative fault (Bray & Rodriguez-Marek, 2004) and it is important as it may influence the size of the displacement for the structures with linear behaviour (Moustafa & Takewaki, 2010). Relevant pulses are those having amplitude over 50% from the peak velocity and their number can be related to the number of asperities in the fault, according to Sommerville (2000).

The phase of the pulse θ_p encompasses information about the shape of the pulse and its damage potential, by indicating the speed of the energy release (Mavroeidis & Papageorgiou, 2003; Vassiliou & Makris, 2011). Similar information is provided by the pulse shape parameters, which reflect the "building" speed of pulses and have an important influence on the magnitude of the structural dynamic response: the more abrupt is the pulse, the higher the shock on the structure, leading to higher peak responses (He & Agrawal, 2008).

The energy content of the pulse is another indicator of the damage potential of the ground motion and is estimated from the cumulative squared velocity/acceleration of the extracted pulse signal or, in the case of wavelet analysis, from the squared absolute wavelet coefficients corresponding to the pulse (Baker, 2007; Zamora & Riddell, 2011). The instant of the pulse arrival differentiates the pulses caused by directivity effects from those having different causes (Baker, 2007).

The elastic and inelastic pseudo-response spectra are a common approach employed for the characterization of seismic ground motions, as the responses of the structures encompass information about the excitations. Since PLGMs significantly affect flexible structures, the velocity and displacement spectra are relevant for this purpose.

3.3.2. Pulse models

Various functions are used either for approximating/extracting the pulse components from PLGMs or for simulation purposes. They are expected to meet some general criteria, which include having simple mathematical expressions and use as few parameters as possible in their formulation and with clear physical meaning (Mavroeidis & Papageorgiou, 2003; He & Agrawal, 2008; Moustafa, 2010). Interestingly, some of these functions also lead to closed-form solutions when deriving the response of single-degree of freedom (SDOF) oscillators (e.g. Mavroeidis et al., 2004).

The models existing in the literature vary from simple waveforms like those presented in Figure 3.4, to oscillatory functions with decaying envelopes (i.e. modified wavelet forms) as those displayed in Figure 3.5.

The simplest shapes used for approximating/extracting the pulses include limited duration harmonic functions and simple rectangular or triangular functions. Makris & Chang (2000a, 2000b) proposed the use of cycloidal pulse shapes with varying number of oscillations.

$$\begin{aligned}
&\textit{Type } A\text{: } \ddot{u}_p(t) = \omega_p \frac{\nu_p}{2} \sin\left(\omega_p t\right) && 0 \le t \le T_p \\
&\textit{Type } B\text{: } \ddot{u}_p(t) = \frac{\nu_p}{\omega_p} - \frac{\nu_p}{\omega_p} \cos\left(\omega_p t\right) && 0 \le t \le T_p \\
&\textit{Type } C_n\text{: } \ddot{u}_p(t) = \omega_p \nu_p \cos\left(\omega_p t + \varphi\right) && 0 \le t \le \left(n + \frac{1}{2} - \frac{\varphi}{\pi}\right) T_p
\end{aligned} \tag{3.1}$$

In Alavi & Krawinkler (2001) triangular pulse shapes are employed for approximating the pulse. Xie et al. (2005) proposed to represent the pulses using successions of simple pulse wave-forms. They studied the effectiveness of eight sets of simple pulse waveforms in characterizing the FD and FS effects, starting from the basic shapes rectangular, sinusoidal, triangular and quadratic (Figure 3.4). The number of simple pulses in the succession depends on the effect (FD/FS) and the shape depends on the

building speed of the pulse (shock loading or gradual loading). The different types of pulses are compared in terms of acceleration/velocity amplitude and SDOF oscillator responses.

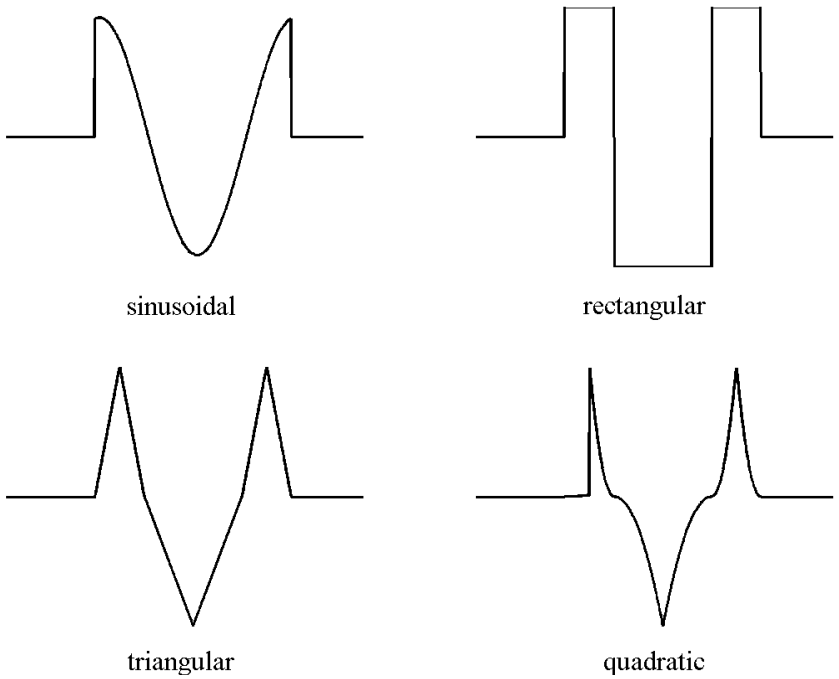

Figure 3.4. Simple waveform pulse models

More sophisticated models consist in diverse types of wavelets employed for the approximation of the velocity pulse. A widely used wavelet, calibrated for simulating motions with no remnant displacement, is the modified Gabor wavelet proposed by Mavroeidis & Papageorgiou (2003). In order to ensure closed-form solutions for the response for the SDOF oscillator excited with this pulse model, the Gaussian envelope in the Gabor wavelet is replaced with an elevated cosine function. The analytical form of the M&P wavelet becomes (see also Figure 3.5):

$$v_p(t)=\begin{cases} A_p\dfrac{1}{2}\left[1+\cos\left(\dfrac{2\pi f_p}{\gamma}\left(t-t_p\right)\right)\right]\cos\left[2\pi f_p\left(t-t_p\right)+\nu\right], & for -\dfrac{\gamma}{2f_p}\le t\le\dfrac{\gamma}{2f_p} \\ 0 & , otherwise \end{cases} \quad (3.2)$$

$$\frac{1}{f_p}<\frac{\gamma}{f_p}, where the\ frequency\ f_p=\frac{1}{T_p}$$

In the previous formula A_p controls the amplitude of the signal, f_p the frequency of the pulse, γ accounts for the number of oscillations, ν for the phase angle and t_p is the instant where the peak amplitude of the pulse is attained. By derivation or integration the corresponding acceleration/displacement trace can be determined. In the case of the displacement pulse, the integration constant is set equal to zero. The pulses are extracted from the records by performing a simultaneous fitting of the acceleration, velocity and displacement time-histories and the corresponding elastic response spectra. The parameters for defining the pulse are correlated with the earthquake characteristics: the pulse frequency f_p is highly dependent on the moment magnitude and type of soil, while the pulse amplitude A_p is related to the seismic wave velocity and to the rise time.

In Dabaghi et al. (2011) the M&P wavelet formulation for the velocity pulse is modified in order to ensure zero final displacement automatically, by doing the following manipulation:

$$v_p(t)=\left[\frac{1}{2}A_p\cos\left(\frac{2\pi\left(t-t_p\right)}{T_p}+\nu\right)-\frac{D_r}{\gamma T_p}\right]\left[1+\cos\left(\frac{2\pi}{\gamma}\frac{(t-t_p)}{T_p}\right)\right],\left(t_p-\frac{\gamma T_p}{2}\right)\le t\le\left(t_p+\frac{\gamma T_p}{2}\right) \quad (3.3)$$

where D_r is the residual displacement and has the following form:

$$D_r=\frac{A_pT_p}{4\pi(1-\gamma^2)}\left[\sin\left(\nu+\gamma\pi\right)-\sin\left(\nu-\gamma\pi\right)\right] \quad (3.4)$$

The Gabor wavelet is proposed for pulse approximation by Dickinson & Gavin (2011):

$$v_p(t)=A_pe^{-\frac{\pi^2}{4}\left(\frac{t-t_p}{\gamma T_p}\right)^2}\cos\left(2\pi\frac{t-t_p}{T_p}-\nu\right) \quad (3.5)$$

A similar definition is proposed by Fu & Menun (2004), i.e.:

$$v_p(t) = \begin{cases} A_p e^{1-\alpha(t-t_p)-e^{-\alpha(t-t_p)}} \sin\left(2\pi/T_p\,(t-t_0)\right), \text{ for } t_0 < t \le 2T_p \\ 0, \quad \text{otherwise} \end{cases} \qquad (3.6)$$

$$\text{where } t_p = t_0 + 0.75T_p$$

The instant of the peak amplitude t_p is defined as a function of the arrival time t_0 and the pulse period. In an earlier publication (Menun & Fu, 2002), the pulse is defined in two steps, using two different shape parameters, one in the time interval from [0, $3/4T_p$] and the second in the interval up to $2T_p$. The calibration of the model is made using field recorded time-histories and their pseudo-velocity response spectra.

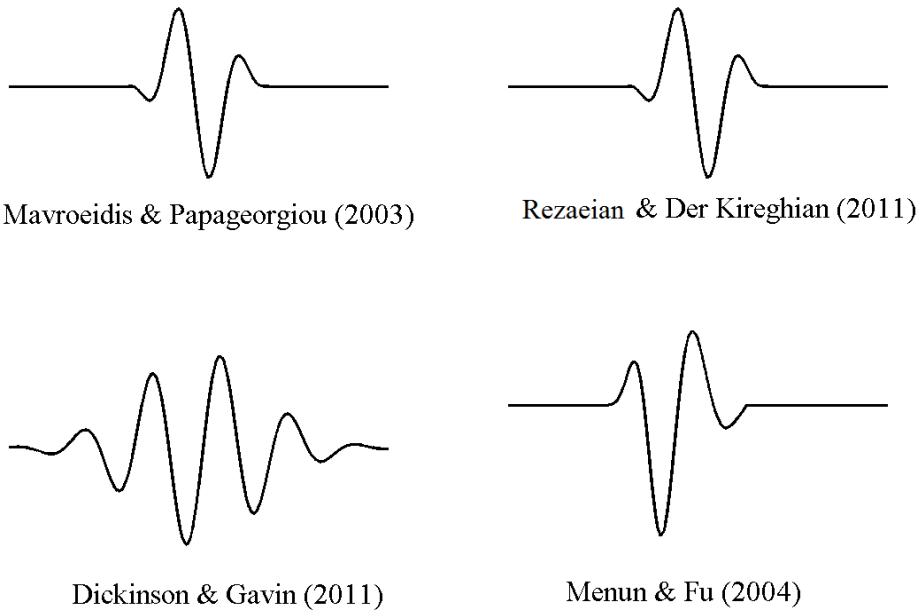

Figure 3.5. Wavelet-based pulse models

He & Agrawal (2008) use the following analytic formulation based on the Berlage wavelet:

$$v_p(t) = Ct^n e^{-at} \sin\left(2\pi f_p t + \nu\right), \qquad t \ge t_0 \qquad (3.7)$$

The parameter t_0 marks the beginning of the pulse, while the decay factor α and the variable n determine its shape. The velocity pulse together with the acceleration and

displacement versions are fitted to a number of recorded time-histories and their corresponding response spectrums are derived in order to demonstrate the performance of the proposed model.

Mukhopadhyay & Gupta (2013a) propose the use of the first and second derivatives of the Gaussian function for modelling velocity pulses, given by the following equations:

$$v_1(t) = Ate^{-\frac{t}{2\sigma^2}} \tag{3.8}$$

$$v_1(t) = A\left(1-\frac{t^2}{\sigma^2}\right)te^{-\frac{t^2}{2\sigma^2}} \tag{3.9}$$

with A controlling the amplitude of the pulse and σ indicating the width of the function.

Moustafa & Takewaki (2010) proposed two alternative models: one deterministic and one stochastic. The velocity pulse is defined by combining a harmonic or a stationary oscillatory part and a time-varying envelope.

$$v_p = A(t)v_{st}(t) \tag{3.10}$$

Two options are provided for modelling the envelope, characterized by the following parameters: A_p controlling the amplitude, the parameters α, β, λ for the shape of the pulse and t_p for the instant of peak velocity.

$$\begin{aligned} A_1(t) &= A_p\left(e^{-\alpha t}-e^{-\beta t}\right),\ \beta>\alpha>0 \\ A_2(t) &= A_p e^{-\left[\frac{2\pi(t-t_p)}{\lambda}\right]^2} \end{aligned} \tag{3.11}$$

A deterministic approach is proposed for modelling the frequency content of the pulse by means of a superposition of n sinusoids with various frequencies f_{pi}, one for each pulse to be modelled, i.e.

$$v_p(t) = \sum_{i=1}^{n} A_{pi} \sin[2\pi f_{pi}(t - t_{pi})], \tag{3.12}$$

By differentiating Equation(3.12), the acceleration pulse(s) is obtained in a closed-form solution. A second, probabilistic model is also proposed. In this case, the amplitudes of the pulses are defined as a set of uncorrelated, zero-mean random variables, satisfying the following conditions:

$$\begin{aligned} &A_{pi}A_{pj} = \sigma_i^2 \delta_{ij} \ , \quad i = 1, \ldots, n \\ &\begin{cases} \delta_{ij} = 1 & for\, i = j \\ \delta_{ij} = 0 & for\, i \neq j \end{cases} \end{aligned} \tag{3.13}$$

where δ_{ij} is the Kronecker delta function and σ_i^2 is the variance of the amplitude. The power spectrum density function describing the low frequency content for the case when n pulses are considered is given by:

$$S_{v_p}(\omega) = 2\sum_{i=1}^{n} \sigma_i^2 \delta\left(\omega - 2\pi f_{pi}\right) \tag{3.14}$$

The stationary samples compatible with the thus defined power spectrum density are modulated using the envelope functions given in equation (3.11). From equation (3.14) the acceleration power spectrum of the stationary part can be obtained as

$$S_{a_p}(\omega) = \omega^2 S_{v_p}(\omega) \tag{3.15}$$

3.4. IDENTIFICATION AND EXTRACTION OF PULSES

3.4.1. Record fitting

The commonly employed approach for the identification and extraction of pulses from PLGM records consists in fitting deterministic pulse-models to the corresponding time-histories (Menun & Fu, 2002; Moustafa & Takewaki, 2010; Yaghmaei-Sabegh, 2010). When this operation is carried out in the time domain, the velocity traces are usually used since they are characterized by narrow frequency content making pulses easier to identify. Alternatively, fitting of the model's response spectrum to that of the recorded time-history has also been considered. Makris & Black (2004a, 2004b) studied the response spectra of the acceleration traces of pulse-type inputs (type A and type B cycloidal pulses) and proposed an energy-based length scale for characterizing this type of motions. This length scale, a function of the pulse acceleration and pulse period ($L_e = a_p T_p^2$) is used to normalize the displacements, leading to a dimensionless response spectrum (Π-response spectrum). By comparing this with the spectrum of a deterministic model, it is easier to identify the existence of a pulse in the ground motion than by inspecting their time-domain representations, since it brings "*a remarkable order in the inelastic response of the structures*" (Vassiliou & Makris, 2011). Later on, based on this work, Tang & Zhang (2011) propose a methodology for establishing the pulse period and amplitude from the congruence between the Π-response spectrum of predefined wave forms and the response spectrum of the velocity and acceleration traces. Mavroeidis & Papageorgiou (2003) performed the record fitting simultaneously in the time and spectral domains.

3.4.2. Time-frequency representation analyses

3.4.2.1. Empirical mode decomposition versus wavelet transform for seismic data analysis

The EMD and the WT are popular TFR techniques used for the identification – extraction of pulses and for the characterization of pulse-like records. As in the case of record fitting, most of the analyses are performed on the velocity time-histories. Comparative studies have been made on the performance of these two techniques in characterizing highly non-stationary signals (see for example Huang et al, 1998; Huang et al, 2001; Kijewski-Correa & Kareem, 2006).

The advantages of the EMD over the WT highlighted by Huang et al (1998) and Huang et al (2001) after analysing records of oceanic wave motions and earthquake data, refer to the high adaptability to the data thanks to the record-dependant timescales and the avoidance of the uncertainty principle limitations.

An alternative perspective of the potential and limitations of these techniques is provided by Kijewski-Correa & Kareem (2006). The higher adaptability to the data of the EMD in comparison with the WT is acknowledged, however its resistance to statistical analyses resulting from this property is also recognized. On the other hand, the rigorous mathematical background of the WT has a more pronounced impact on the resulting signal representations. It is concluded that any of these techniques are suitable for analysis purposes as long as the user is aware of their limitations.

3.4.2.2. The empirical mode decomposition for pulse identification

The EMD has been employed for the analysis of specific acceleration or velocity time-histories, proving its capability to identify and isolate the specific pulse. Loh et al (2001) used the EMD for accelerogram decomposition and identified and isolated the

pulse by visually assessing the energy contribution of each IMF to the total energy. The IMFs containing higher frequencies represent the "dynamic part", while the remaining low-frequency IMFs are classified as the pulse part. In Zhang et al (2005) the authors process the velocity time-histories and separate the pulse by choosing the IMFs which appear to contain the pulse. Xu & Agrawal (2010) use the EMD to characterize the broadband part of the record remaining after the pulse is extracted by fitting a deterministic function. The resulting IMFs are grouped depending on their frequency content in relation to the pulse frequency, resulting in a higher frequency component and lower frequency component, whose impact on the structural behaviour is inspected.

Overall it can be observed that due to its empirical formulation, a relevant amount of subjectivity is involved when using the EMD for the characterization of pulse-like records. The records are decomposed in a small number of components, but their number varies from one record to another and cannot be known prior to the analysis. Although the number of IMFs can be pre-set to a fixed value, this is not recommended, since this defies the purpose of using the EMD in the first place (adaptability to the data) and it might end-up causing mode-mixing or creating unrealistic effects in the IMFs (Huang et al, 2003; Rilling et al., 2003). A second observation is the fact that the frequency content of the IMFs differs from one record to another. Although the frequency content decreases as the number of the IMF increases, it cannot be established what frequency content each IMF, or a group of IMFs, will encompass. This becomes a challenge when trying to analyse a database containing several records. To overcome this, Zhang et al. (2003) establish ranges of frequencies incorporated by each IMF, after visually observing a number of decomposed records, and band-pass filter them based on this in order to facilitate the analyses.

On the other hand, due to the great adaptability of the algorithm to the data, the information carried by individual/groups of IMFs has physical meaning, matching well

with different parts of the total response spectra of the analysed record. Consequently the EMD is mostly used for characterization of individual records rather than for databases.

3.4.2.3. The wavelet transform for pulse identification

The WT has the potential to accurately approximate dominant features of seismic signals by using a reduced number of coefficients. Todorovska et al. (2009) assessed the quality of such approximations, using 1÷8% of the CWT coefficients, in terms of energy, peak power or time of collapse for non-linear oscillator for the velocity traces of strong ground motions and conclude that even a 1% approximation offers good results. The WT has thus been employed as a step in the following type of methodology used for pulse identification and extraction: the WT is applied to the signal, the maximum coefficient is identified and the corresponding wavelet represents the first approximation of the pulse. This procedure is repeated several times, until a satisfactory representation of the pulse is obtained. This algorithm was used by Baker (2007) for extracting the pulses from a database of velocity records. In order to classify them, Baker laid the bases of the first quantitative classification for PLGMs. The pulses were compared with the residual motion (remnant after the extraction of the pulse) in terms of peak ground velocity (PGV) and energy. Based on this information, a pulse indicator was defined for separating PLGMs from non-PLGMs. For an even more precise classification, the pulses caused by FD effects were identified by placing an additional constraint regarding the time of the pulse arrival. The wavelet used by Baker for the analysis (Db 4) was chosen based on visual inspection of the velocity time-histories.

In a later study, Yaghmaei-Sabegh (2010) studies the influence of the type of mother wavelet on the features of the extracted pulses. Two groups of functions are compared: symmetrical bi-orthogonal wavelets (Rbio2.4 and Bior1.3) and unsymmetrical

orthogonal functions (Db4, Db7, Sym4, Coif2 and Haar). The pulses extracted using each type of wavelet are compared in terms of period, shape and pulse indicator, as defined by Baker (2007). It is observed that symmetrical functions preserve the linear phase, while the unsymmetrical ones often cause phase distortions or even exhibit a false pulse preceding the main one. The study is performed on two velocity traces belonging to ground motions classified as being of pulse-type. Although limited by the reduced number of time-histories analysed, the results draw attention to the impact of various types of analysing function on the results; in the end the use of symmetrical, smooth wavelets is recommended by the author for obtaining accurate representation.

A similar study was performed by Vassiliou & Makris (2011) who evaluated the performance of a different group of waves, namely the type B, C1 and C2 cycloidal pulses, the symmetric and the anti-symmetric Ricker wavelets, the Gabor wavelet and the Mavroeidis & Papageorgiou (2003) (M&P) wavelet on extracting the pulses from accelerograms. The influence of the wavelet phase and of the number of oscillations on the outlook of the extracted pulses is investigated, showing that unsymmetrical functions may cause phase distortions of the data. The importance of the smoothness of the function is reasserted, confirming the observation that smoother functions lead better approximations using a reduced number of wavelets (Todorovska et al., 2009). In order to improve the adaptability to the data, 2 degrees of freedom are incorporated in the definition of the Gabor wavelet and M&P wavelets, allowing for the variation of phase and number of oscillations aside from translation and dilation of the function, thus indirectly proposing the chirplet transform (Mann & Haykin, 1991; Spanos et al. 2007) as an alternative technique for the analysis of PLGMs.

Mollaioli & Bosi (2012) applied the WT to the energy time-histories corresponding to the signals, rather than to the velocity/acceleration traces. The absolute and relative input energies are defined as functions of the earthquake recorded velocity/acceleration

and the structural responses. The peak wavelet coefficients indicate the time-frequency position of the largest amount of energy which corresponds to the pulse.

3.4.3. Low-pass filtering

A different approach was employed by Ghahari et al. (2010) for extracting pulses from a collection of PLGMs. The authors used a moving average (MA) filter to separate the pulse from the higher frequency content. The cut-off frequency of the filter is defined as a function of the main pulse period, which is approximated by applying several times the STFT: the window size is gradually increased until it incorporates the dominant pulse of the velocity time-history.

In Mukhopadhyay & Gupta (2013a) a pulse indicator is employed first in order to check if the motion qualifies as pulse-type. This is based on the largest amount of energy between two successive zero-crossings encountered in the velocity time-history. The time-location of the velocity pulse is also identified based on the same energy-between-zero-crossings principle. By "smoothening" the accelerogram in the pulse time-window identified (which represents essentially a type of low-pass filtering) the pulse is isolated and extracted from the accelerogram. Based on the inspection of the pulses extracted from the time-histories, two pulse-like functions are proposed for the modelling of the low frequency content (see Eqs. (3.8) & (3.9)).

3.5. MODELLING OF PULSE-LIKE GROUND MOTIONS

The need for artificial time-histories comes from the limited amount of records available in various regions or for specific site conditions. This is of particular concern to PLGMs as a limited number of records exist, associated with a handful of seismic events (e.g. Baker, 2007; Mukhopadhyay & Gupta, 2013). In this manner the need to account for the aleatory uncertainty in seismic risk assessments can be addressed. Indeed, artificial

accelerograms are used in engineering applications for investigating and improving the structural design and for the probabilistic assessment of seismic demands (e.g. Taflanidis, 2011; Taflanidis & Jia, 2011; Taflanidis & Vetter, 2011).

There are two types of earthquake models which are used for simulation purposes – seismological models and phenomenological (or "record-based") models (e.g. Vetter & Taflanidis, 2014). The first type consists of complex models which incorporate information about the source-to-site geometry, fault characteristics, soil conditions, seismic waves' arrival and frequency content among others. The latter are simpler, empirical models, which focus less on the realistic representation of the mechanisms causing the earthquake and more on reproducing structural responses similar to those observed for field recorded accelerograms. They are defined based on the features characterizing in the time-histories of seismic records, usually regardless of the causative mechanisms.

With regards to PLGMs, a limited and relatively small number of records are identified as being of pulse-type, with a significant percentage of them coming from the proximity of San Andreas Fault. Due to these limitations, their modelling is a challenging task. Consequently, most of the models are either fully phenomenological or defined as a combination between a seismological model for the higher-frequency content and a phenomenological model for the pulse part.

Since a significant amount of the total energy is released through the pulse (Moustafa & Takewaki, 2010), PLGMs are sometimes represented by means of only simple, deterministic waveforms (see Section 3.3.2). However, this type of representation does not account for the entire range of frequencies which may affect structures (including the higher modes of flexible structures). Therefore, in order to obtain more realistic simulations, the higher frequency (HF) content needs to be also taken into account (Sommerville, 1998; Makris & Chang, 2000a; Mavroeidis & Papageorgiou, 2003;

Mukhopadhyay & Gupta, 2013b). As a result, the following formulation is commonly adopted for PLGM representation:

$$y_{PL}(t) = y_P(t) + y_{HF}(t) \tag{3.16}$$

The pulse part y_P in equation (3.16) is usually obtained through differentiation from the velocity models previously presented in Section 3.3.2, while for the high-frequency content y_{HF} probabilistic models proposed for the modelling of regular (pulse-free) ground motions can be employed. Empirical relationships between the parameters of the functions and various characteristics of the earthquake (source properties, rupture velocity, magnitude, soil conditions, epicentral distance etc.) are obtained by means of regression analyses on the databases of recorded PLGMs considered (Mavroeidis & Papageorgiou, 2003; Mavroeidis et al., 2004; Baker, 2007; Mukhopadhyay & Gupta, 2013b among others). In some cases, in order to account for the variability observed in the pulse properties, probability distributions have been estimated for the pulse model parameters (Dabaghi et al., 2011; Dickinson & Gavin, 2011; Gavin & Dickinson, 2011; Taflanidis & Jia, 2011).

Various alternative models are used in the literature for modelling the HF part of the record. Mavroeidis & Papageorgiou (2003) use the specific barrier model which consists in defining the frequency content as a summation of small sub-events with delayed arrival times and characterized by different frequency content and amplitude, as it is usually the case of seismic waves.

Fu & Menun (2004) define the HF content as a superposition of cosines with random phases (Φ_k) and varying amplitudes (B_k). The temporal variation is accounted for by means of a time-varying envelope (A_n).

$$s_{hf}(t) = A_n(t)\sum_{k=1}^{n} B_k \cos(\omega_k t + \Phi_k),\ k = 1,\ldots,N$$
$$A_n(t) = e^{1-\beta(t-t_p)-e^{-\beta(t-t_p)}}\sqrt{-\beta+\beta e^{\left(-\beta(t-t_p)\right)^2}+(2\pi/T_p)^2} \tag{3.17}$$

In the previous equation, t_p is the instant of peak amplitude, β controls the shape and T_p is the pulse period. According to the authors, the instant when the peak of the pulse and the peak of the higher frequency content occur are strongly related.

Dickinson & Gavin (2011) propose the spectral representation method to generate time-histories compatible with a Kanai-Tajimi spectrum, which are then modulated using a time-varying envelope. The envelope is a function of 4 parameters which define its evolution over time.

$$a_{hf}(t) = A(t;\tau_0,\tau_1,\tau_2,\tau_3)\sum_{k=1}^{N+1}\sqrt{S(f_k,f_g,\zeta_g)}\cos(2\pi f_k t_i + \theta_k),$$
where
$$A(t;\tau_0,\tau_1,\tau_2,\tau_3) = \begin{cases} ((t_i-\tau_0)/\tau_1)^2 & for\ \tau_0 \le t \le \tau_0+\tau_1 \\ 1 & for\ \ \tau_0+\tau_1 \le t \le \tau_0+\tau_1+\tau_2 \\ e^{-(t_i-\tau_0-\tau_1-\tau_2)/\tau_3)} & for\ \tau_0+\tau_1+\tau_2 \le t \end{cases} \tag{3.18}$$
$$S(f_k,f_g,\zeta_g) = \frac{(2\zeta_g f_k/f_g)^2}{(1-(f_k/f_g)^2)^2+(2\zeta_g f_k/f_g)^2}$$

Dabaghi et al. (2011) consider amplitude-modulated white noise filtered through a time-varying filter to represent the residual acceleration. The white noise is first processed using a filter whose characteristics (central frequency and frequency bandwidth) vary in time. Next, a time-varying envelope is used to modulate the amplitude as in:

$$s_{hf}(t) = A(t)\left\{\frac{1}{\sigma_h(t)}\int_{-\infty}^{t} h\left[t-\tau,\lambda(\tau)\right]w(\tau)d\tau\right\} \tag{3.19}$$

The parameter $\lambda(\tau)$ controls the time-variation of the filter, depending on the frequency and damping, $\sigma_h(t)$ normalizes the variance of the integral to 1 and $A(t)$ is the time-varying envelope.

The seismological stochastic model developed by Boore (2003) for generating frequency content higher than 0.1Hz is employed by Taflanidis (2011) and Taflanidis & Jia (2011) for modelling the higher frequency content. The power spectrum shape is defined as a combination between the contribution of the source of the earthquake (E), the path effects (P) and the site conditions (G) and the type of motion (I):

$$S(\omega | M_W, R) = E(M_W, \omega) P(R, \omega) G(\omega) I(\omega) \tag{3.20}$$

In the previous equation M_W represents the moment magnitude and R the distance from the source. The duration of the simulations are controlled by means of a time-varying envelope given by the equation

$$a(t) = a(t/t_p)^b e^{-c(t/t_p)} \tag{3.21}$$

where the parameters a, b and c control the shape of the envelope and t_p corresponds to the instant of the peak.

A modelling approach alternative to the superposition of low and high frequency content is proposed by Yamamoto & Baker (2013) for the simulation of fault normal accelerograms. For this purpose, a database of records is analysed by means of the MWPT and the corresponding power spectral density is estimated. The resulting WT coefficients are separated based on their energy content in two groups – the major group, containing few coefficients which contribute to most of the total energy (70%) and the minor group containing a large number of coefficients which account for the remaining energy content (30%). The probability distribution functions to be used for

modelling the position and the evolution of the energy content on the time-frequency plane are established by means of statistical analyses. A total number of 13 parameters are needed to completely define the model. The resulting model accounts for the entire frequency content of the accelerogram at once, leading to an integrated model of the high and low frequency parts. However, since it relies on the wavelet transform it carries along some of the challenges experienced by this technique when dealing with very low frequency content, leading to large standard deviations for long period spectra. For this reason, the authors suggest the use of the proposed model in conjunction with techniques more suitable for simulating low-frequency pulses.

CHAPTER 4 : PROPOSED MODEL FOR PULSE-LIKE GROUND MOTIONS

4.1. PRELIMINARY REMARKS

The commonly used approach for the stochastic modelling of PLGMs is by superposing two models characterizing complementary frequency content, according to Eq. (3.16). The low frequency content y_P is usually represented using analytically defined (velocity) functions, which are further differentiated in order to obtain the pulse part for the synthetic accelerogram. The variability in the pulse characteristics is taken into account through probability distributions obtained by analysing collections of records. Still their reliability in accounting for the uncertainty is limited by the scarcity of existing records which have been classified as being pulse-type. An alternative, probabilistic approach for the definition of pulses was proposed by Moustafa & Takewaki (2010) and consists in defining their spectral content by means of delta functions (see Eq. (3.12) – (3.15)).

The higher frequency content y_{HF} is modelled using non-stationary stochastic models which account for the evolutionary behaviour either in time or in both time and frequency (see Section 3.5 and Appendix B).

Herein a novel non-separable non-stationary stochastic PLGM model is proposed, motivated by the phenomenological observation that pulses appear as low-frequency patches of energy on the time-frequency plane. This observation is made after analysing several PLGMs from a time-frequency perspective by using signal processing techniques reviewed in Chapter 2. Further, two new stochastic models for the pulses are defined. For the higher frequency content existing stochastic models commonly used in

civil engineering applications are employed. Finally, the methodology for generating pulse-like accelerograms is also presented.

4.2. PULSE-LIKE RECORDS: A SIGNAL PROCESSING PERSPECTIVE

In order to have a reliable phenomenological model for PLGMs it is important to appropriately define the signature features of recorded PLGMs, namely the long period, high amplitude oscillations. Focusing on this aspect, it has been argued in the literature that the pulses encountered in PL records can be separated in "acceleration" and "velocity" pulses (Vassiliu & Makris, 2011; Zamora & Riddell, 2011; Mollaioli & Bosi, 2012). In the first category fall the pulses visible in the accelerogram which are known to be causing extreme damage; in the second one, pulses which are not visible in accelerograms, but only in the velocity and displacement time-histories. The latter category is assumed to be the result of a non-zero-mean *"succession of high-frequency one-sided acceleration spikes"* (Makris & Black, 2004c).

From a modelling perspective, the case of the "acceleration" pulses is clear since pulses are readily visible in all the time-histories, thus they can be taken into account using the modelling approach given by Eq. (3.16). In what follows, the case of several records characterized by "velocity" pulses is further investigated from a signal processing viewpoint to clarify the modelling approach for such cases.

In particular, accelerograms belonging to the following earthquakes – 1971 San Fernando (California), 1989 Loma Prieta (California), 1992 Landers (California) and 1999 Kocaeli (Turkey) are being processed using the GHWT, MWPT and ST (presented in Sections 2.4 - 2.5). The acceleration signal corresponding to the 0÷5 rad/sec frequency interval is reconstructed and integrated to obtain the corresponding

velocity trace. This procedure is equivalent to a LP filtering of the accelerogram, using a 5rad/sec cut-off frequency.

According to the definition given by Makris & Black (2004c) for the "velocity" pulses, it is expected that once the high-frequency content is removed, the PGV will significantly decrease, i.e. the pulse will not appear anymore in the corresponding velocity traces. In Figure 4.1 and Figure 4.2 the filtered accelerograms are superposed to the original time-histories. The acceleration pulse extracted by Baker (2007) is also plotted for reference and referred to as the Db4 CWT. In the right hand side panels, the "velocity" signals obtained by integrating the filtered accelerograms are superposed on the velocities corresponding to the recorded accelerogram. The correlation coefficients between the actual velocity traces and those corresponding to the filtered accelerogram are also listed in Table 4.1.

Table 4.1. Correlation coefficients between the ground motion velocity and the velocity corresponding to the filtered accelerograms

	1989 Loma Prieta St. Saratoga	**1992 Landers St. Lucerne**	**1999 Kocaeli St. Gebze**	**1971 San Fernando St. Pacoima Dam**
ST	**0.93**	**0.98**	**0.98**	**0.78**
GHWT	**0.92**	**0.98**	**0.98**	**0.80**
MWPT	**0.97**	**0.99**	**0.99**	**0.91**
Db4 CWT	**0.64**	**0.79**	**0.77**	**0.69**

The results clearly show that very small differences exist between the two velocity traces and usually consist in some very low-amplitude higher frequency noise. Therefore, what is referred in the literature as "velocity" pulses represents the low-frequency content already present in the acceleration trace. The integration performed in order to obtain the velocity trace only improves its visibility by zooming-in in the existing low-frequency content. This is further clarified in Appendix A, where it is

shown that the integration is equivalent to a low-pass filtering of signals (Worden, 1990).

The separation of pulses in "acceleration" and "velocity" pulses is therefore not justified. A less ambiguous way to refer to pulses would be as the low-frequency content of the recorded accelerograms. Further, if a classification from the perspective of their visibility due to the amount of higher frequency content characterizing accelerograms is to be made, an appropriate way (at least from a signal analysis viewpoint) would be by considering the so-called signal-to-noise ratio, where the noise corresponds to the high-frequency content of the accelerogram.

A second comparative analysis is performed on the previously considered record of Landers earthquake and the less "noisy" record (according to the above interpretation) belonging to the 1999 Imperial Valley (California) earthquake. The pulses approximated by Mavroeidis & Papageorgiou (2003) are superimposed on the accelerogram and velocity time-histories in order to facilitate the visualisation of the low-frequency content (Figure 4.3, Figure 4.4). The time-frequency representations of the acceleration and velocity traces for each case are obtained by alternatively employing the wavelet-based techniques presented in Chapter 2. By investigating the frequency content between 0÷10 rad/sec of the accelerogram and velocity traces, the previous observation regarding the low-frequency content of the accelerogram is confirmed. Low-frequency energy patches, located in similar positions appear in the TF representation of both, regardless of how visible the pulses are in the accelerogram. In fact, the pulse is not visible in the Landers record because of the very low signal-to-noise ratio, i.e. the high frequency components are very intensive.

Based on these analyses, it is reasonable to introduce the pulses directly into the accelerograms, without necessarily considering the velocity traces. Accordingly, in the following section a non-separable non-stationary stochastic model for simulating pulse-

like accelerograms is proposed, which uses the same stochastic modelling approach for representing both the HF and the LF energy content. The model offers great versatility in modelling the HF content, being able to accommodate various representations for this purpose (e.g. spectrum compatible, phenomenological or seismological models), while the proposed approach for generating the pulses represents a more versatile alternative to other stochastic models existing in the literature (Moustafa & Takewaki, 2010).

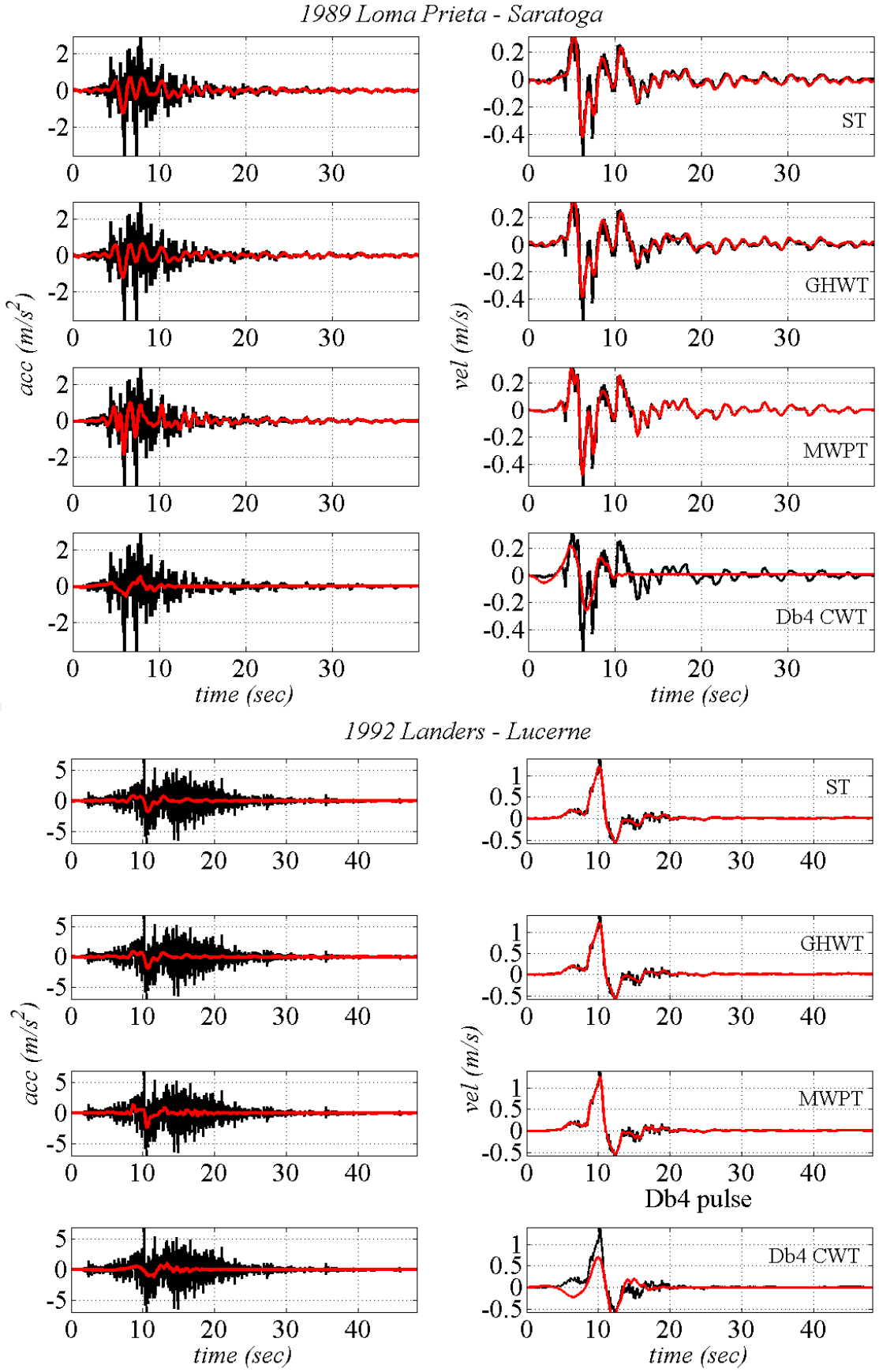

Figure 4.1. Filtered accelerograms and the corresponding velocity traces obtained (I)

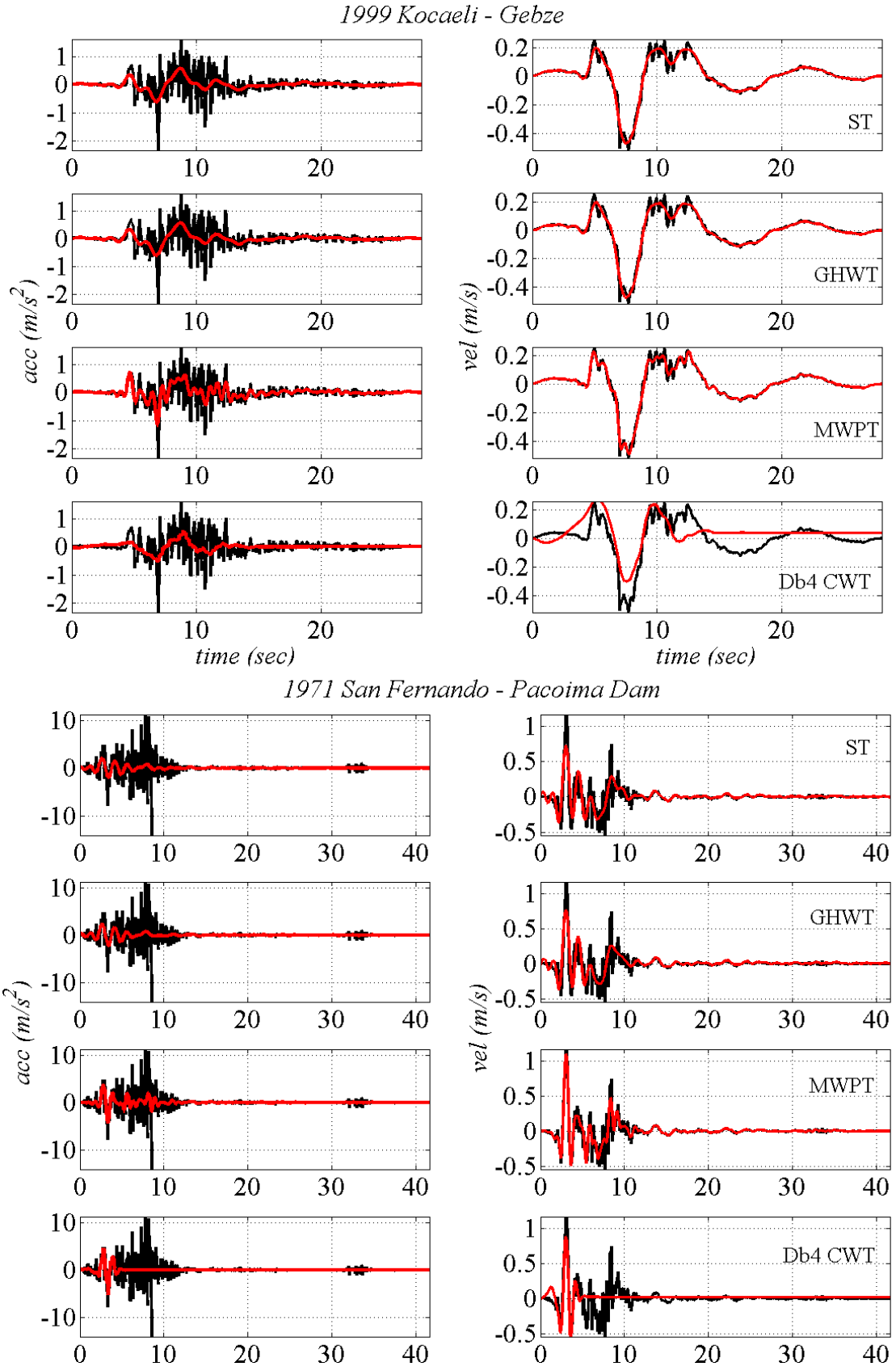

Figure 4.2. Filtered accelerograms and the corresponding velocity traces (II)

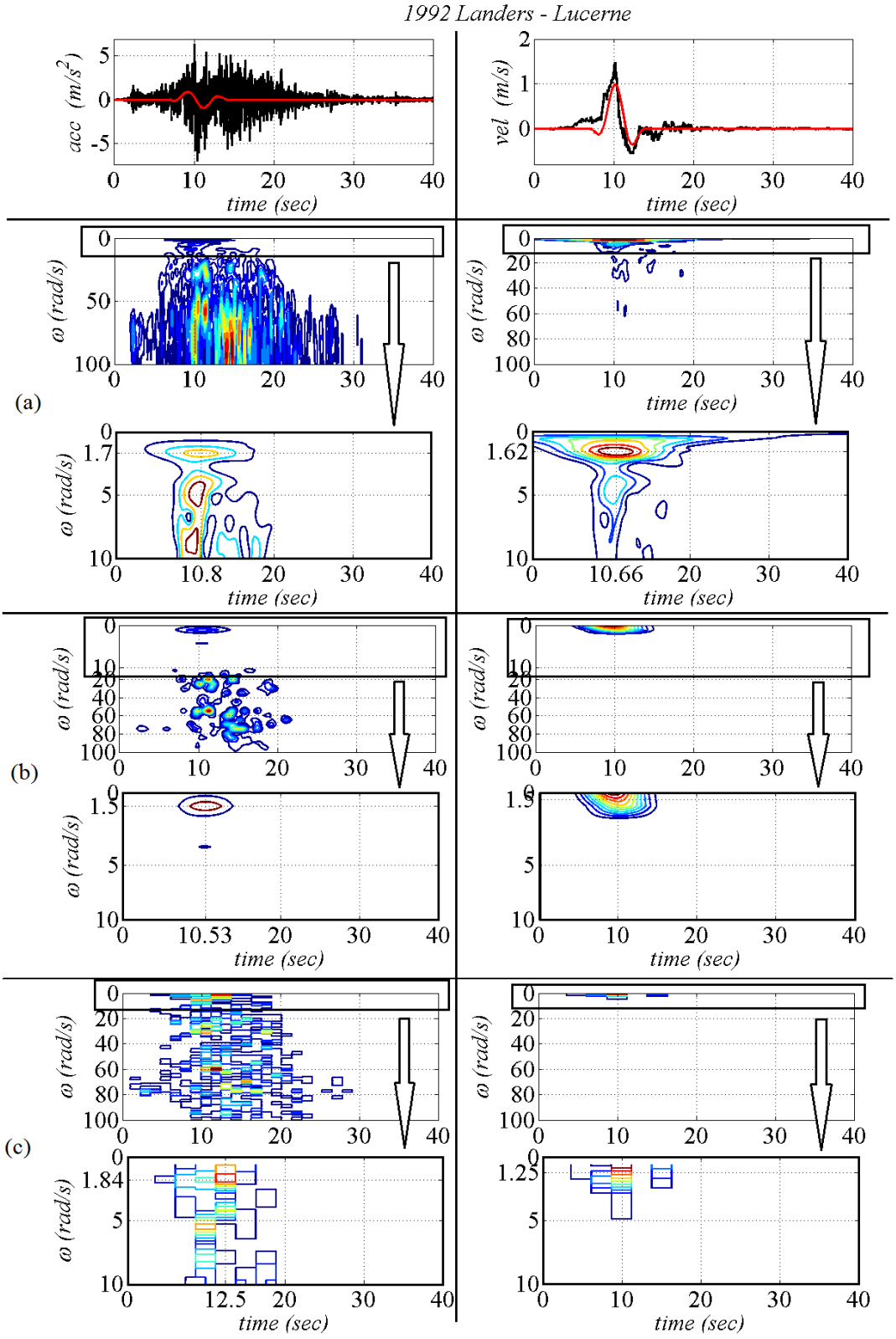

Figure 4.3. Time-frequency representations for an accelerogram extremely corrupted with higher frequency components: (a) S-transform; (b) harmonic wavelet transform; (c) Meyer wavelet packets transform

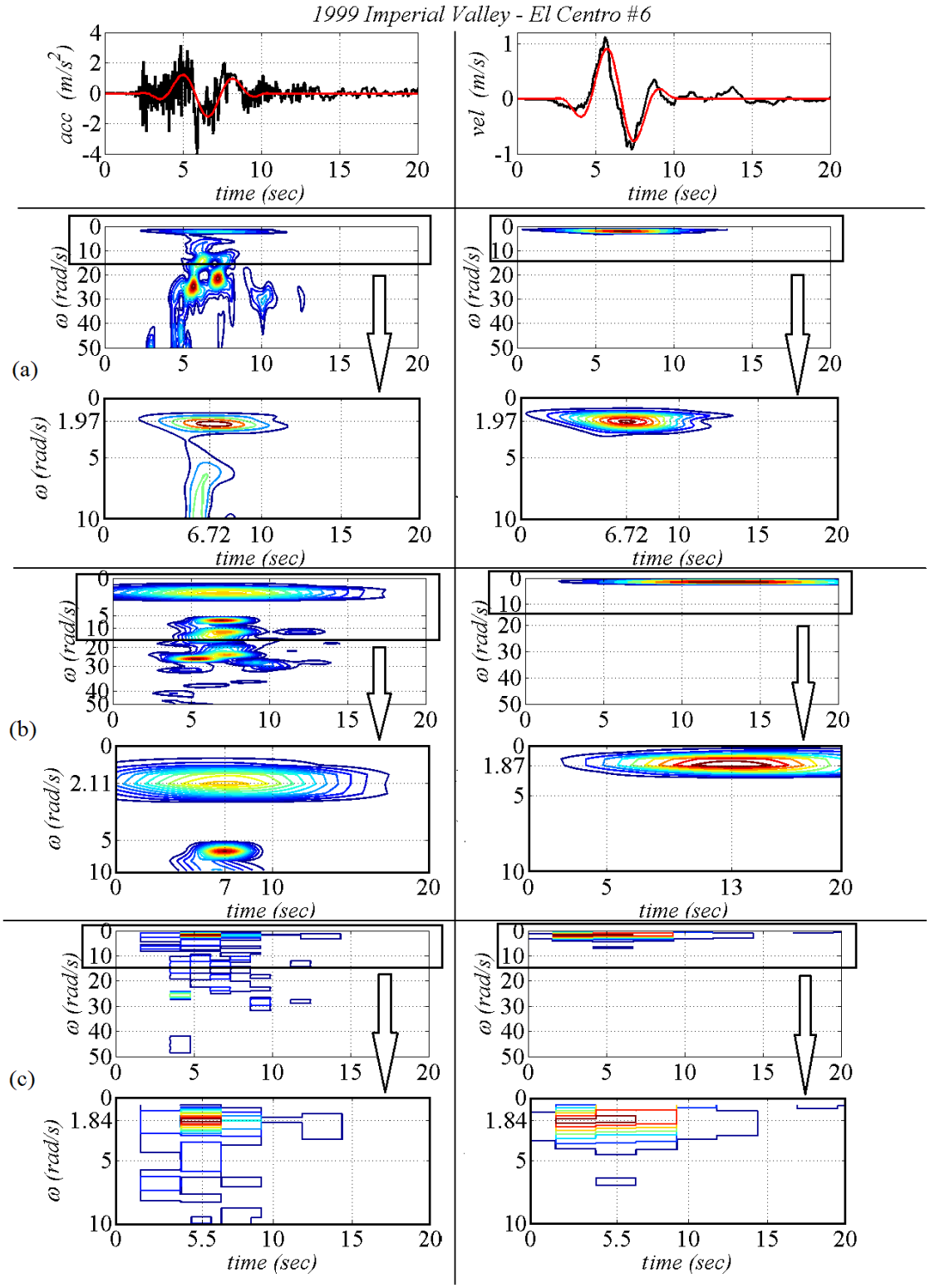

Figure 4.4. Time-frequency representations for an accelerogram reasonably corrupted with higher frequency components: (a) S-transform; (b) harmonic wavelet transform; (c) Meyer wavelet packets transform

4.3. FULLY STOCHASTIC MODEL FOR PULSE-LIKE EARTHQUAKES

Based on the approach used for modelling PLGMs as a superposition of low and high frequency (LF/HF) components given by Eq. (3.16), consider the following definition for a PLGM model:

$$y_{PL}(t)=\sum_{i=1}^{P} a_{HF_i}(t)g_{HF_i}(t)+\sum_{j=1}^{R} a_{LF_j}(t)g_{LF_j}(t),\ P,R\geq 1 \tag{4.1}$$

In the previous equation, adjacent frequency bands characterizing the frequency content of PL seismic signals are modelled by means of separable uncorrelated processes. Each process is defined as a product between a stationary zero-mean random process *g(t)*, characterized by power spectrum density *G(ω)* and an envelope function *a(t)* with an evolutionary behaviour over time (see Appendix B for further details). It has been previously shown in the literature that the energy distribution of amplitude modulated processes can be represented using the concept of evolutionary power spectrum density (EPSD) under the assumption that the envelope functions *a(t)* vary sufficiently slow in time (Priestley, 1965; Conte & Peng, 1997; Spanos & Failla, 2004), i.e.

$$EPSD(t,\omega)=|a(t)|^2 G(\omega) \tag{4.2}$$

Consequently, the EPSD characterizing the energy distribution of the pulse-process given in Eq. (4.1) on the time-frequency plane reads as

$$EPSD_{PL}(t,\omega)=\sum_{i=1}^{P}\left|a_{HF_i}(t)\right|^2 G_{HFi}(\omega)+\sum_{j=1}^{R}\left|a_{LF_j}(t)\right|^2 G_{LF_j}(\omega) \tag{4.3}$$

Considering the simplest possible case in which the low-frequency pulse and the higher frequency content can be modelled using singular uniformly modulated processes for each part ($P = R = 1$), the model simplifies to (Lungu & Giaralis, 2013).

$$y_{PL}(t) = a_{LF}(t)g_{LF}(t) + a_{HF}(t)g_{HF}(t) \tag{4.4}$$

$$EPSD_{PL}(t,\omega) = EPSD_{LF}(t,\omega) + EPSD_{HF}(t,\omega) \tag{4.5}$$

as shown in Figure 4.5.

For most cases the above model is considered to be sufficient for modelling the frequency content of PLGMs (Section 3.5). However, in case of recorded accelerograms characterized by richer low-frequency content (e.g. Mukhopadhyay & Gupta, 2013a), more than one process can be considered for representing the pulse part, as it is illustrated later on in Section 5.3. Further, a more detailed representation of the higher frequency content using the superposition of several amplitude modulated processes can be adopted (e.g. Conte & Peng, 1997).

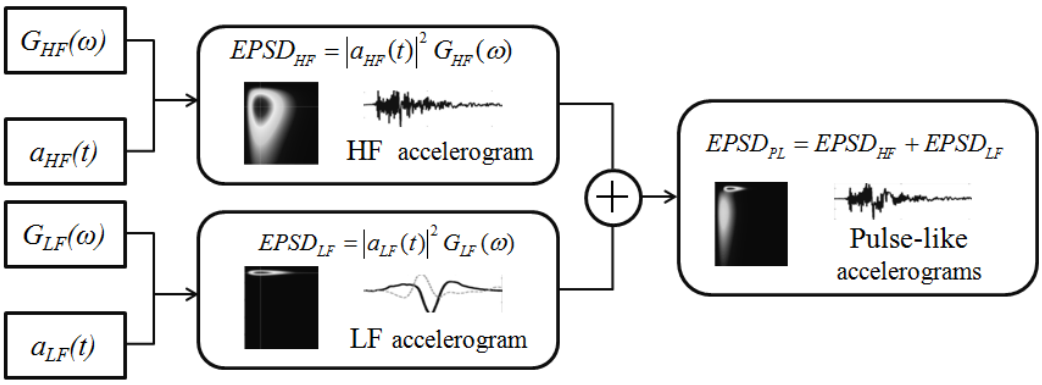

Figure 4.5. Application of the proposed non-separable non-stationary stochastic pulse-like ground motion model for the special case of $P = R = 1$

In fact, stochastic processes of this form, with various analytical expressions for the envelope function a_r and the power spectrum G_r, have been used in the literature to model the earthquake induced strong ground motion in terms of acceleration for various

earthquake engineering applications. For instance, Spanos & Vargas-Loli (1985) have considered this type of stochastic model for the generation of spectrum compatible accelerograms in a stochastic framework. Conte & Peng (1997) have used a similar model for the characterization and representation of specific field recorded accelerograms associated with specific historic seismic events, while Spanos & Failla (2004) have used it for assessing the efficiency of various wavelet bases in estimating the underlying evolutionary power spectrum of non-stationary processes.

4.4. HIGH FREQUENCY PROCESS MODELLING

4.4.1. Power spectrum density G_{HF}

Herein, the Clough-Penzien (CP) spectral form (Clough & Penzien, 1993) which is commonly used in civil engineering applications is considered for modelling the HF content (Figure 4.6). This power spectrum density is obtained by filtering Gaussian white noise using a CP filter, which is given by the equation:

$$G_{CP}(\omega)=\frac{\left(\frac{\omega}{\omega_f}\right)^4}{\left(1-\left(\frac{\omega}{\omega_f}\right)^2\right)^2+4\zeta_f^2\left(\frac{\omega}{\omega_f}\right)^2}\times\frac{1+4\zeta_g^2\left(\frac{\omega}{\omega_g}\right)^2}{\left(1-\left(\frac{\omega}{\omega_g}\right)^2\right)^2+4\zeta_g^2\left(\frac{\omega}{\omega_g}\right)^2},\ \omega\le|\omega_{\max}| \quad (4.6)$$

The CP filter is obtained as a succession of two filters: a high pass filter, characterized by the cut-off frequency ω_f and the steepness (slope of the filter) ζ_f and the Kanai-Tajimi filter (Kanai, 1957), which accounts for the soil conditions in a simplified way (soil stiffness ω_g and damping ratio ζ_g). The role of the high pass filter is to eliminate the low-frequency components allowed by the Kanai Tajimi filter, which are the source

of exaggerate responses in the case of long period structures. The term ω_{max} represents the maximum frequency allowed in the spectrum (i.e. cut-off frequency).

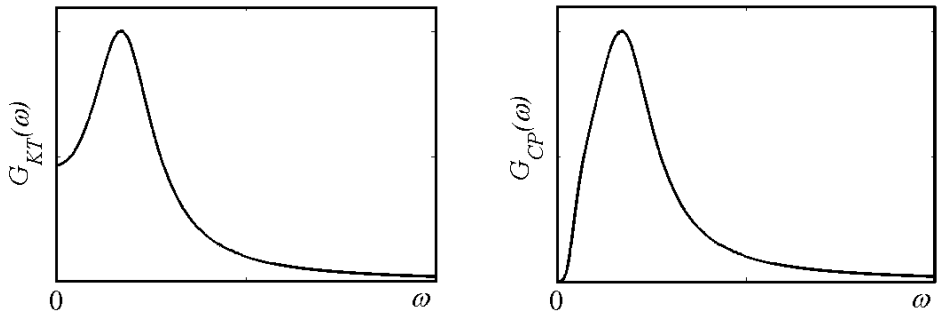

Figure 4.6. Power spectra shapes for representing the broadband frequency content: Kanai Tajimi (left) and Clough Penzien (right)

Alternatively, other options exist for modelling the broadband frequency content of the PLGM accelerograms, either phenomenological or seismological. A review of the commonly used models for ordinary accelerograms is provided by Shinozuka & Deodatis (1988), while a retrospective of the approaches used in the case of PLGMs is presented in Section 3.5.

4.4.2. Time-varying envelope a_{HF}

The evolution over time of the ground motion intensity is determined by means of an envelope function. Some commonly used shapes can be found in (Shinozuka & Deodatis, 1988). Herein the bell-shaped function proposed by Bogdanoff et al (1961) and given by the following equation is considered:

$$a(t) = Cte^{-\frac{bt}{2}} \tag{4.7}$$

The parameter C is proportional with the peak amplitude of the envelope, while b controls the width and is related to the effective duration T_{eff} of the ground motion. The envelope can be defined employing the non-linear relationship between T_{eff} and the

parameter *b* developed by Giaralis & Spanos (2012) for the case when the effective duration is defined as the time-interval in which 90% of the total energy is released (Trifunac & Brady, 1975), i.e.:

$$T_{eff} = t_{95} - t_{05} \tag{4.8}$$

The terms t_{05} and t_{95} in the previous equation represent the time instants at which 5% and 95% of the total energy is cumulated.

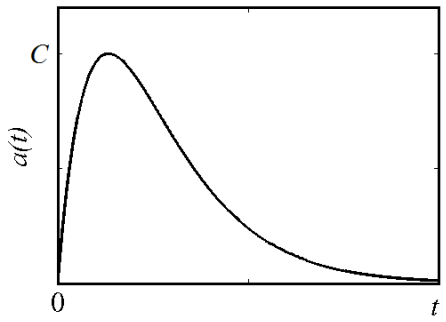

Figure 4.7. The BGB envelope function

4.5. PROPOSED LOW-FREQUENCY PROCESS MODELLING

4.5.1. Power spectrum density G_{LF}

Based on the appearance and characteristics of the low-frequency content (Section 3.3, Section 4.2), two new spectral shapes are proposed herein for the definition of the LF process stationary part (Figure 4.8): a box-like shape (BOX) and a raised cosine shape (COS).

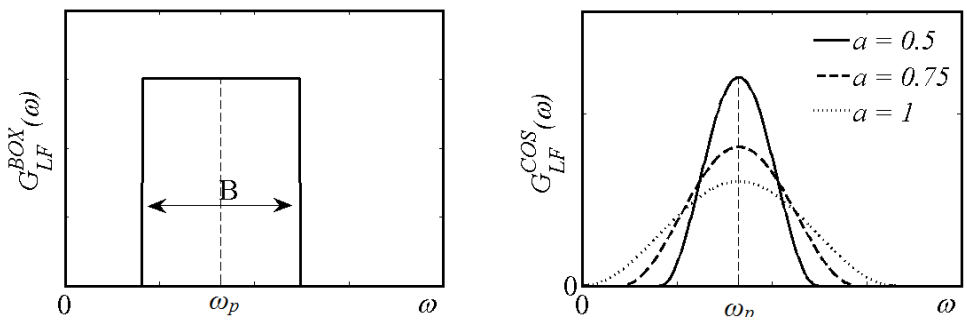

Figure 4.8. Power spectrum density shapes for the pulse content: BOX (left) and COS (right)

The BOX shape is given by the following equation

$$G_{LF}^{\mathrm{BOX}}(\omega)=\begin{cases}1,\ for\ \omega_p - B/2 \le \omega \le \omega_p + B/2\\ 0,\ otherwise\end{cases} \tag{4.9}$$

where the term B represents the bandwidth of the low-frequency content characterizing the pulse and ω_p is the dominant frequency of the pulse. Due to its simplicity this shape can be easily employed in conjunction with random vibrations analyses techniques incorporating stochastic processes expressed by means of the harmonic wavelet transform (Spanos & Kougioumtzoglou, 2012).

The equation for the COS spectrum is

$$G_{LF}^{COS}(\omega)=\frac{1}{2}\left(1+\cos\left(\frac{2\pi}{2\alpha\omega_p}\left(\omega-\omega_p\right)\right)\right),\quad for\ \ \omega(1-\alpha)\le\omega\le\omega(1+\alpha) \tag{4.10}$$

where $\alpha \le 1$ is a shape parameter. This second shape is more appropriate for time-histories simulation thanks to its smoothness (e.g. Taflanidis & Jia, 2011). Similarly to the previous case, the power spectrum is centred at the pulse dominant frequency ω_p.

4.5.2. Time-varying envelope a_{LF}

The following envelop function depicted in Figure 4.9 is proposed herein for accounting for the evolutionary behaviour of the pulses:

$$a(t)=C_{LF}e^{-\frac{1}{2}\left(\frac{\omega_p}{\gamma}\left(t-t_p\right)\right)^2} \tag{4.11}$$

In the previous formulation, C_{LF} controls the amplitude, ω_p represents the dominant frequency of the pulse, t_p is the instant of the peak and the parameter γ controls the shape of the envelope. Similar envelopes have been used by Tian et al. (2007) and Moustafa & Takewaki (2010) for this purpose.

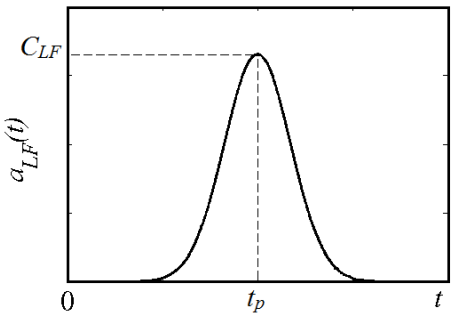

Figure 4.9. Time-varying envelope for the pulse part

4.6. GUIDELINES FOR SIMULATING PULSE-LIKE ACCELEROGRAMS

Pulse-like accelerograms can be readily generated using the herein proposed model by taking the following steps:

I. Definition of the higher-frequency content

- Establish the level of detail for the higher frequency content, i.e. the number of sub-processes to be employed for its representation;
- For each composing sub-process, set the parameters C and b for the time-varying envelope and ω_g, ζ_g, ω_f and ζ_f for the Clough-Penzien spectrum (see Section 4.4). Obtain the corresponding functions by using Eq. (4.6) and (4.7);
- Combine each power spectrum with the corresponding time-varying envelope according to Eq. (4.2) to obtain the amplitude-modulated power spectra densities characterizing each sub-process;

II. Definition of the low-frequency content

- Establish complexity of the low-frequency content, i.e. the number of pulses;

- For each pulse, set the parameters for the envelope C_{LF}, t_0 and γ and for the power spectrum set ω_p and B for the BOX shape or, alternatively, ω_p and α for the COS shape (see Section 4.5).
- Combine each power spectrum with the corresponding time-varying envelope according to Eq. (4.2) to obtain the amplitude-modulated power spectra densities characterizing each pulse;

III. Generation of high-frequency part (HF samples) and of the pulse part (LF samples) of the accelerogram

- Generate samples compatible with each sub-process using appropriate simulation techniques. Any qualified random field generation method for stationary random processes (e.g. the spectral representation method or the autoregressive-moving average method presented in Appendix C) can be employed for the purpose.

IV. Generation of pulse-like accelerograms

- Superpose the HF samples with the corresponding LF samples in order to obtain pulse-like accelerograms. The resulting non-separable, non-stationary time-histories are characterized by the evolutionary power spectrum previously defined in Eq. (4.5).

The methodology for simulating artificial pulse-like accelerograms compatible with the herein proposed model, for the special case when a total number of two processes are considered is exemplified in Figure 4.10. A total number of 11 parameters are needed to completely define the pulse-like process: 6 for the definition of the higher frequency content and 5 for the definition of the pulses.

By employing the herein proposed representation for modelling and simulation purposes, some of the previously mentioned limitations of the existing models can be avoided. The simple definition and the flexibility allows for the use of the model for Monte Carlo simulations (Taflanidis & Jia, 2011) or alternatively, as input for various

stochastic dynamic techniques such as statistical linearization, stochastic averaging etc. (Spanos & Kougioutzoglou, 2012).

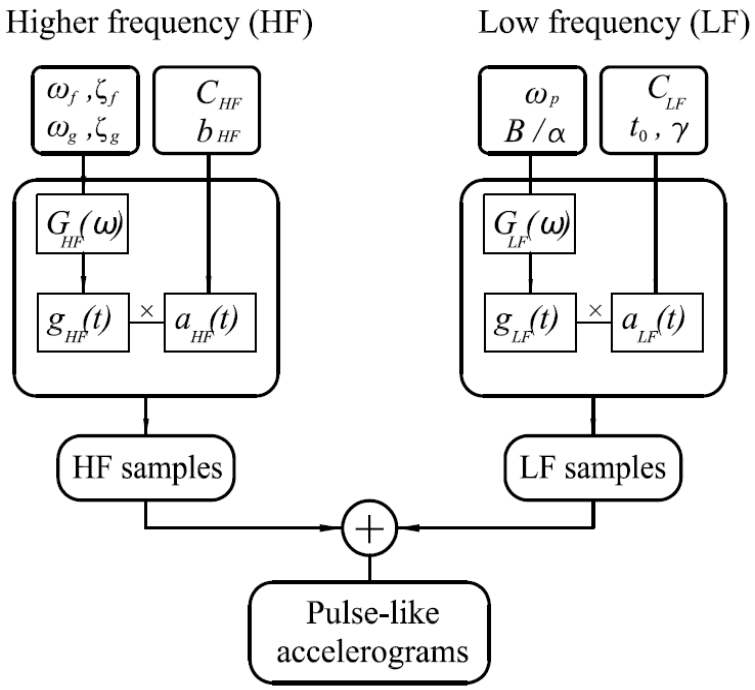

Figure 4.10. Methodology for generating artificial pulse-like accelerogram for the special case of *P* = *R* = *1* in Eq. (4.4)

CHAPTER 5 : USE OF THE PROPOSED PULSE-LIKE GROUND MOTION MODEL

5.1. PRELIMINARY REMARKS

In this chapter the potential of the stochastic model previously introduced to capture salient features of recorded PLGMs is assessed. First, the capability of the low-frequency models (Section 4.5) to realistically capture characteristics of pulses extracted from field recorded time-histories is investigated. For this purpose, a database of acceleration pulses from the literature is considered and for each entry equivalent synthetic pulse processes are defined. Samples compatible with each process are generated and the average characteristics are compared with the target values. Since the model proposed herein is a record-based one, the aim is to obtain similar structural responses as in the case of the extracted pulses, thus a comprehensive verification from the structural response perspective is performed. Furthermore, empirical predictive equations for the low-frequency models parameters are obtained through regression analysis against earthquake magnitudes and distances from the fault. These relationships are used later on in Chapter 6 to generate pulse processes for specific seismic scenarios. Secondly, the parameters for defining a complete PLGM process (i.e. which incorporates both low frequency and high frequency content) characterizing a specific field recorded accelerogram are determined. A simple, record based representation is chosen for modelling the higher frequency content, i.e. by means of modulated Clough-Penzien spectrum, while for the low-frequency content the two alternative shapes are considered. It is demonstrated that the proposed model is capable to realistically

simulate the structural impact of the considered record, while at the same time accounting for the variability in its features.

Finally, the potential of the PLGM model to generate ensembles of code-compatible accelerograms which include pulse effects is also illustrated.

5.2. FIT OF THE LOW-FREQUENCY MODELS TO A GIVEN DATABASE OF PULSES

5.2.1. Description of the database

A set of 91 acceleration *pulses* is considered herein for the calibration of the proposed low-frequency models (Appendix D). They have been extracted from a subset of accelerograms belonging to the Next Generation Attenuation ground motion library. These have been classified as pulse-type by Baker (2007) using a weighting scheme for the approximated pulse with regards to the original time-history. The characteristic pulses are approximated after performing the wavelet-decomposition of the velocity traces, as the superposition of the ten Db4 wavelets (Figure 5.1) which result into the highest coefficients. The acceleration pulses which will be used further in this work are obtained by numerically differentiating the thus obtained velocity pulses and are provided in http://www.stanford.edu/~bakerjw/pulse-classification_old.html (accessed 25-03-2011).

The selection of the database to be used for calibration is arbitrary. It should be kept in mind that the quality of the low-frequency content approximation depends on the algorithm used for extraction and may vary across a database. If an alternative database of pulses, i.e. obtained by using a different weighting scheme or a different wavelet function (e.g. Vassiliou & Makris, 2011; Zamora & Riddell, 2011 or Mukhopadhyay &

Gupta, 2013b) is employed for the calibration, different sets of pulse parameters may be obtained.

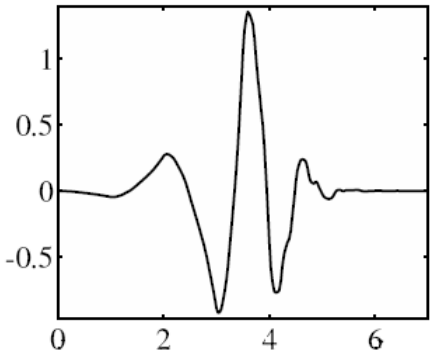

Figure 5.1 Daubechies wavelet function of order 4 (Db4) used by Baker (2007) for pulse extraction

5.2.2. Preliminary values for the parameters

In order to calibrate the low-frequency models against the database considered, the following methodology is adopted: initial choices for the parameters are established and employed for defining pulse processes corresponding to each entry in the database. Next, ensembles of accelerograms compatible with each process thus defined are generated to assess the capacity of the pulse models to replicate, on average, the characteristics of the pulses in the database. Based on these initial results, the parameters of the model which need further adjustment are identified and discussed. Finally, regression analyses are performed in order to establish connections between the model parameters and the seismological characteristics for the herein considered database.

The following vectors of parameters, specifying the dominant pulse frequency ω_p, the bandwidth α/B, the peak amplitude C_{LF}, together with its location t_p and its shape γ_{LF}, need to be defined for each entry in the database:

$$\begin{aligned} LF_{COS} &= [\omega_p,\ \alpha,\ C_{LF}, \gamma_{LF},\ t_p] \\ LF_{BOX} &= [\omega_p,\ B,\ C_{LF}, \gamma_{LF},\ t_p] \end{aligned} \tag{5.1}$$

The dominant frequency ω_p is derived from the pulse periods T_p estimated by Baker (2007) and repeated herein in Appendix D, using the formula:

$$\omega_p = \frac{2\pi}{T_p^{Baker}} \tag{5.2}$$

In order to establish a suitable choice for the bandwidth of the pulse for the case of the BOX definition (Eq. (4.9)), the power spectral densities of the pulses in the database are estimated as the square value of the FT coefficients. The bandwidth is defined as the frequency interval around the main peak which includes the power spectral density values higher than 10% of the peak value. The correspondence between the resulting bandwidths, the pulse period and the number of records can be seen in the histogram plots in Figure 5.2. Since there is a linear correspondence between the bandwidth estimated this way and the pulse frequency, the parameter B is defined as

$$B = \omega_p \tag{5.3}$$

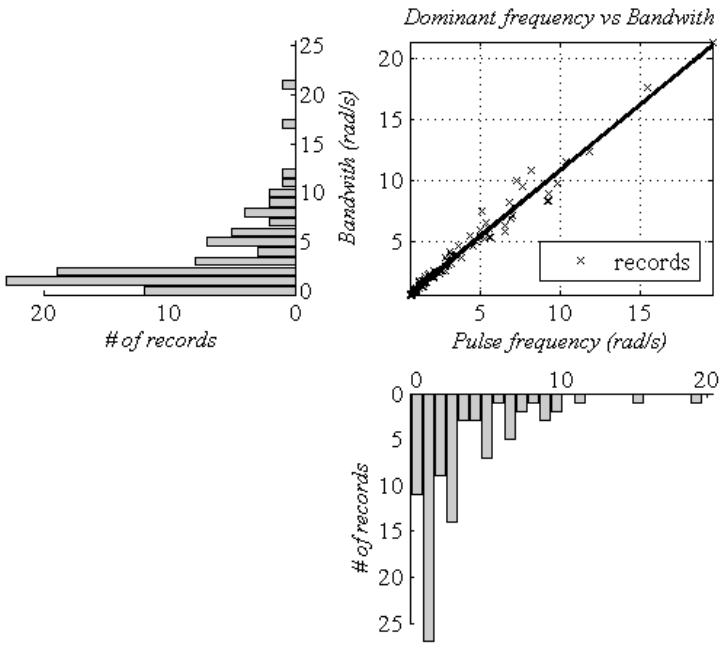

Figure 5.2. Histograms of the bandwidth and dominant frequencies of the pulses across the database considered

With regards to the shape parameter α used for the definition of the COS spectrum, a slight dependence on the pulse period T_p is observed, i.e. the value decreases as the pulse frequency lowers. However since no clear connection was found, a constant value of 0.5 is adopted throughout this study, which leads to an overall acceptable fit.

The time-varying envelope is calibrated using the methodology employed in Stafford et al. (2009) for defining an "energy-based" envelope function for the simulation of pulse-free accelerograms. This approach is considered suitable since it accounts quantitatively for the energy content and its distribution, ensuring realistic structural responses and reliable synthetic accelerograms according to the authors. Note that the same approach was used by Dickinson & Gavin (2011) who employed it for calibrating the envelope proposed for modelling the higher frequency content which remains after the pulse is removed.

In the first step, the envelopes of the acceleration pulses are evaluated using the following two-steps approach (Dugundji, 1958): the pre-envelope of the pulse y_P is determined as

$$z(t) = y_P(t) + iH\left(y_P(t)\right) \tag{5.4}$$

where the operator $H(\cdot)$ accounts for the Hilbert transform of the considered signal.

Next, the envelope is estimated as the magnitude of the analytic signal defined as

$$a(t) = \sqrt{y_P^2(t) + H^2\left(y_P(t)\right)} \tag{5.5}$$

Envelopes estimated using this approach for several pulses in the database are presented in Figure 5.3.

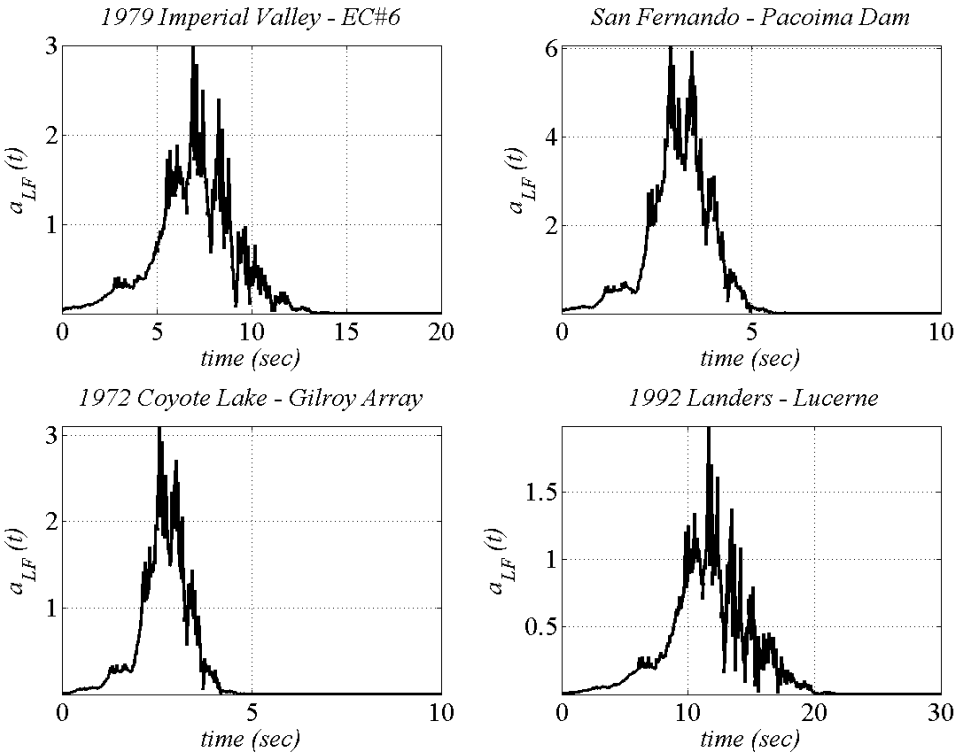

Figure 5.3. Envelopes estimated for pulses extracted from field recorded accelerograms

By recalling Parseval's theorem regarding the conservation of power when shifting between time and frequency domains (Eq. (2.3)) and following the approach used in Stafford et al. (2009), the formulation for the envelope function is obtained as given in the following equation:

$$\int_{-\infty}^{+\infty} a^2(t)\,dt = \int_{-\infty}^{+\infty} y_P^2(t)\,dt + \int_{-\infty}^{+\infty} H^2\left(y_P(t)\right)dt$$

$$\int_{-\infty}^{+\infty} a^2(t)\,dt = 2\int_{-\infty}^{+\infty} y_P^2(t)\,dt$$

$$a(t) = \sqrt{2y_P^2(t)} \tag{5.6}$$

Further, the Arias Intensity I_A and the Husid plot $H(t)$ of the acceleration pulses $y_P(t)$ are also calculated in order to determine the amount of energy and its accumulation over time:

$$I_A = \frac{\pi}{2g}\int_0^{\infty} y_P^2(t)\,dt \tag{5.7}$$

$$H(t) = \frac{\int_0^{t} y_P^2(t)\,dt}{\int_0^{\infty} y_P^2(t)\,dt} = \frac{\pi}{2gI_A}\int_0^{t} y_P^2(t)\,dt \tag{5.8}$$

The normalized Husid function $h(t)$ is obtained by taking the time derivative of the Husid plot and gives the shape of the envelope, i.e.

$$h(t) = \frac{dH}{dt} = \frac{\pi}{2gI_A}\,y^2(t) \tag{5.9}$$

Based on this, equation (5.6) becomes

$$a(t) = \sqrt{2\frac{2gI_A}{\pi}h(t)} \tag{5.10}$$

The exponential function proposed by Tian et al. (2007) and by Moustafa & Takewaki (2010) offers a realistic approximation for the shape of the envelopes estimated, and was thus adopted for the function $h(t)$:

$$h(t) = e^{-\left(\frac{\omega_p}{\gamma}(t-t_p)\right)^2} \tag{5.11}$$

The final shape of the envelope (see Figure 5.4), which is given in section (4.4.2) and repeated here for convenience, is obtained as

$$a(t)=\sqrt{\frac{4gI_A}{\pi}e^{-\left(\frac{\omega_p}{\gamma}(t-t_p)\right)^2}}$$

This can be also written as

$$a(t)=C_{LF}e^{-\frac{1}{2}\left(\frac{\omega_p}{\gamma}(t-t_p)\right)^2} \tag{5.12}$$

The envelope's amplitude C_{LF} can be thus related to the to the energy content of the pulse, which is expressed in this case in terms of the Arias Intensity I_A *of the pulse*. Empirical relationships for estimating the I_A for future records based on seismological characteristics have been derived in the literature; however, the performance of these formulas for records incorporating forward directivity effects is limited (Travasarou et al., 2003). Furthermore, the energy carried by the low-frequency pulse represents only a fraction of the total energy in the accelerogram. Although the contribution of the pulse part to the total energy content in terms of I_A is investigated (e.g. Zamora & Riddell, 2011), there is still limited information available in this sense. For these reasons the connection between the I_A of the pulse and the peak amplitude is not exploited in the context of this work.

The parameter γ in the definition of the envelope is attributed either the role of controlling the shape (Tian et al., 2007; Moustafa & Takewaki, 2010) or that of ensuring the frequency content of the envelope is lower than that of the oscillatory part (Mavroeidis & Papageorgiou, 2003). In Figure 5.4 analytic envelope shapes for various values of γ values are plotted. In the left panel the frequency domain representations for each case are superposed to the approximated envelope. Any possible correlation

between this parameter and shape characteristics of the envelope (e.g. rise time, decay time, effective duration) have been investigated; however no systematic relationship could be established.

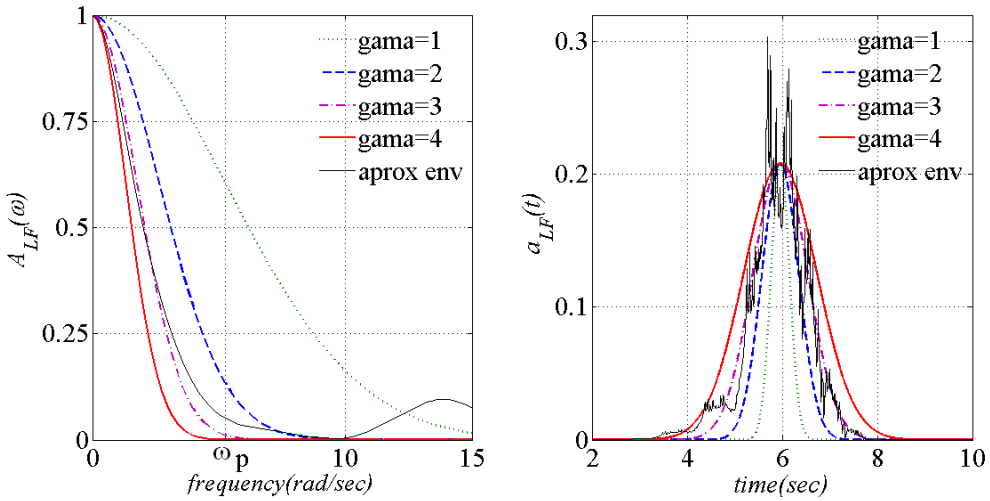

Figure 5.4. Analytic envelopes for various values of γ. Left panel: frequency domain representation (the pulse frequency ω_p is indicated). Right panel: time domain representation.

Eventually, the parameters C_{LF}, γ and t_p are estimated by fitting the function a_{LF} to the approximated envelopes, by using standard nonlinear least squares regression and considering the parameter ω_p as being a record dependant variable, set to the value previously discussed. The fitting is performed using the standard built-in command within the Curve Fitting Toolbox of Matlab.

In order to assess the performance of the model, ensembles of time-histories compatible with each pulse process are generated using the spectral representation method (Appendix C). Accordingly, the synthetic pulses are obtained in two steps: first, stationary samples compatible with $G_{LF}(\omega)$ spectra are generated as a superposition of cosines with random phases Φn, uniformly distributed in the $[0,2\pi]$ interval. Further, each sample is multiplied with a time-varying envelope. For the purpose of this study, a number of 300 samples compatible with each process are generated. Each time history has a duration of 40s and is sampled at a time step $\Delta t = 0.01$s which allows for a maximum frequency to be captured of $\omega_{max} = 2\pi/2\Delta t = 314$ rad/s (50 Hz). Any spurious

very low-frequency or DC components (i.e. non-zero mean value of the signal) are removed by baseline correction of the samples in order to avoid unrealistic velocity/displacement values (e.g. non-zero velocity at the end of the record). An acausal Butterworth high-pass filter with a cut-off frequency of 0.63 rad/s (0.1 Hz), which performs a forward/backward filtering and thus it does not modify on the phase of the signal, is used for this purpose (see also Appendix A).

The aim of the proposed models is to generate realistic structural responses. Consequently, in order to obtain an appreciation of the quality of matching across the database considered, the average 5% elastic pseudo-response spectra for each pulse process are estimated and compared with those corresponding to the original pulses (referred to as the "target" value). In Figure 5.5 the values of the following ratios are plotted against the period of the pulses for the database considered:

$$\rho_{PSA}=\frac{E\left[\int_0^{T_{\max}} S_a^{sim}(T)\,dT\right]}{\int_0^{T_{\max}} S_a^{rec}(T)\,dT};\ \rho_{PSV}=\frac{E\left[\int_0^{T_{\max}} S_v^{sim}(T)\,dT\right]}{\int_0^{T_{\max}} S_v^{rec}(T)\,dT};\ \rho_{D}=\frac{E\left[\int_0^{T_{\max}} S_d^{sim}(T)\,dT\right]}{\int_0^{T_{\max}} S_d^{rec}(T)\,dT} \tag{5.13}$$

A variable quality of matching which depends significantly on the pulse period T_p, is observed across the database. Since an ideal fit would lead to values of the ρ ratios equal to 1, these preliminary results on the performance of the model can be summarized as follows:

- Short pulses - $T_p < 4\sec\ \left(\omega_p > 1.6 rad/\sec\right)$

The response spectra are significantly higher than the target ones, especially for ρ_D. The quality of matching deteriorates as the pulse period reduces, the performance of the model appearing very limited for periods under 1.5sec (hatched areas in Figure 5.5). However, the model proposed herein is targeted for the simulation of low-frequency

components. To this end, pulses with periods less than 1.5s do not qualify as such content.

- Pulses with $4\sec \le T_p \le 6\sec$ $\left(1.1 rad/\sec < \omega_p < 1.6 rad/\sec\right)$

The response spectra are well approximated.

- Pulses with $T_p \ge 6\sec$ $\left(\omega_p < 1.1 rad/\sec\right)$

The response spectra are slightly underestimated compared to the target values.

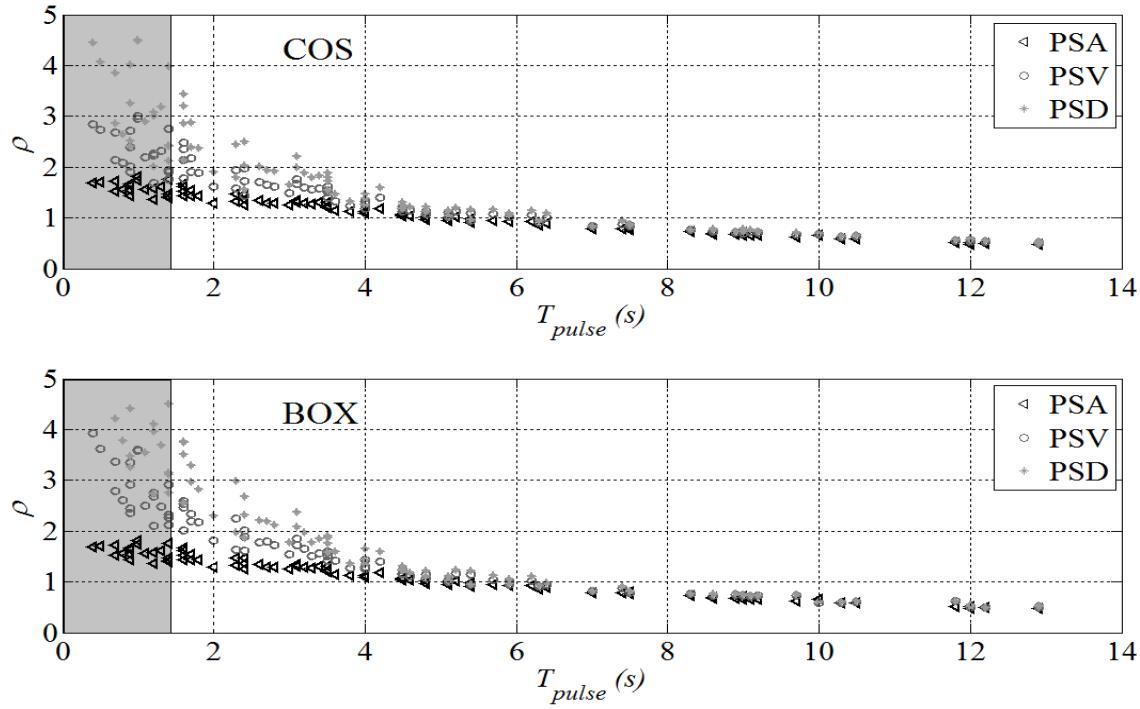

Figure 5.5. Initial quality of matching of spectral ordinates across the database

It should be noted that while the criteria employed for the calibration of the model against the considered database accounts for phenomenological considerations (i.e. envelope fitting), the assessment is made from a structural response perspective. Consequently, based on the preliminary observations regarding the quality of matching in terms of structural behaviour, together with visual investigation of the results, the following aspects are further investigated: the definition of the pulse dominant frequency and the baseline correction applied to the resulting samples.

5.2.3. Definition of the dominant pulse frequency

The periods of the pulses extracted from the considered records are estimated by Baker (2007) based on the wavelet analysis of the velocity traces corresponding to the records considered. The first 10 highest coefficients of the transform are identified and the pulse is expressed as the superposition of the corresponding wavelets. The dominant period corresponding to the thus isolated pulse is defined as the pseudo-frequency of the wavelet (Eqn. (2.6)) corresponding to the highest coefficient, out of the ten used for the pulse reconstruction.

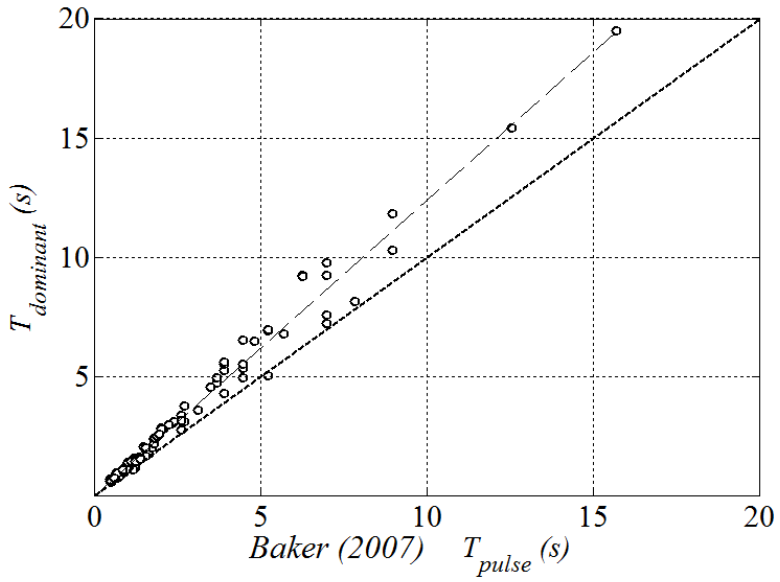

Figure 5.6. Comparison between alternative pulse period definitions (see also Appendix D)

The squared Fourier amplitude spectra coefficients of the acceleration pulses are computed and the frequency ω_p^{peak} corresponding to the peak values of the spectra is identified. The corresponding periods are plotted in Figure 5.6 against the pulse dominant periods derived by Baker (2007) in order to examine the correspondence between the two definitions. Overall it is observed that the pulse period identified by Baker is consistently with at least 20% below the dominant period corresponding to the peak frequency characterizing their power spectra. Subsequently, in the ensuing

calculations thus, the pulse frequency ω_p is taken as the value corresponding to the peak of the acceleration pulse Fourier spectral amplitude, i.e. $\omega_p = \omega_p^{peak}$.

5.2.4. Pulse period dependent baseline adjustment

Similar to field recorded accelerograms, artificial time-histories may need baseline adjustment to remove spurious low frequency or DC components which lead to unrealistic velocity/displacement traces or overestimate structural responses for long period structures (Conte & Peng, 1997; Karabalis et al, 2000; Mukherjee & Gupta, 2002; Boore & Akkar, 2003; Boore & Bommer, 2005). While the source of noise corrupting field recorded accelerograms is either the instrumentation used or the human error (e.g. digitization of analogue records), in the case of artificial accelerograms this has to do with the assumptions made for the purpose of modelling and simulation. The approach used herein for obtaining time-histories consists in generating zero-mean stationary samples compatible with a given power spectrum, followed by a modulation of their amplitude using a time-varying envelope. As explained in Safak & Boore, (1988), the amplitude modulation can induce very low-frequency errors which need to be removed by means of baseline corrections. As previously mentioned, the corrections are performed herein by filtering the samples using an acausal Butterworth high-pass filter with a cut-off frequency of 0.63 rad/s (0.1 Hz), value commonly used for this purpose in the literature (e.g. Conte & Peng, 1997).

Since generally the pulses are characterized by very low frequency content, i.e. the maximum period identified by Baker for example is of 12.9s, which translates into a frequency of 0.49 rad/s (0.08 Hz), careful consideration needs to be paid to baseline corrections. While a common cut-off frequency when processing non-pulse accelerograms using high-pass filters is 0.63rad/s (0.1Hz), such baseline corrections can remove some of the pulse content in the case of pulse-like time-histories and lead to

underestimated structural responses. On the other hand, in the case of shorter pulses a too low cut-off frequency may allow for too much low-frequency content in the simulated time-histories and thus lead to over-estimated responses. This observation is confirmed by the "trend" observed in the preliminary estimation of the quality of fitting shown in Figure 5.5.

For this reason, pulse-dependent baseline corrections are applied to the simulated time-histories. Three period intervals are considered for this purpose:

- for short pulses ($T_p < 4\text{sec}$) the cut-off frequency of the filter is set to 0.2Hz = 1.25 rad/s;
- for medium pulses ($4\text{sec} \le T_p \le 6\text{sec}$) the cut-off frequency of the filter remains set to 0.1Hz = 0.63 rad/s;
- for long pulses ($T_p > 6\text{sec}$) the cut-off frequency of the filter is set to 0.075Hz = 0.47 rad/s.

In Figure 5.7, Figure 5.8 and Figure 5.9 the positive impact of the period dependant baseline corrections on the mean response spectra for each category of pulses can be visualized.

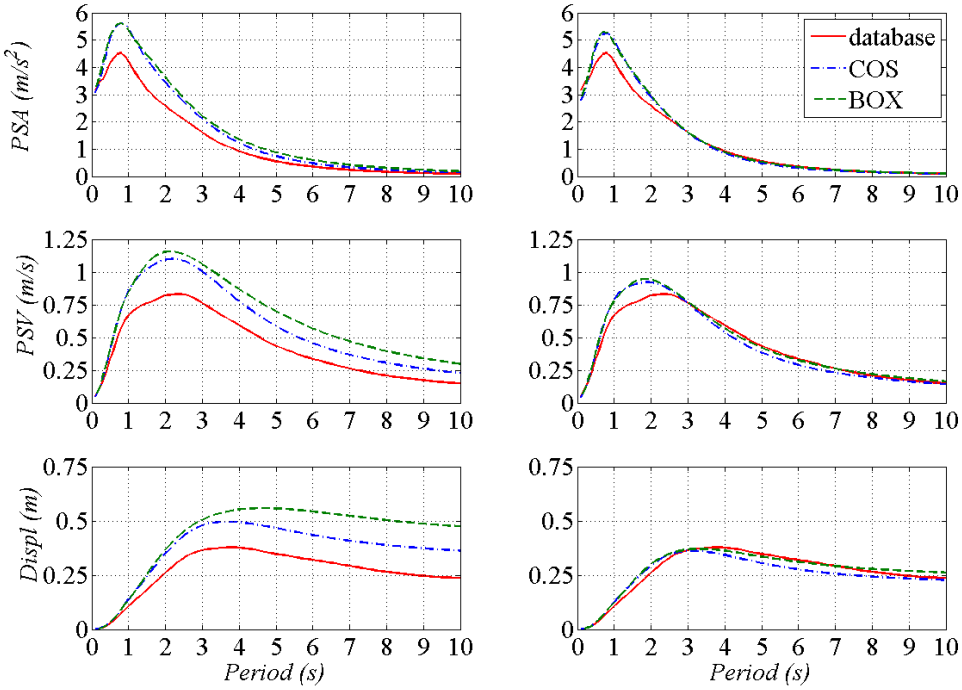

Figure 5.7. Impact baseline corrections on the response spectra – average match for shot pulses (periods under 4s): (left panel) response spectra when using a uniform baseline corrections; (right panel) response spectra when using a pulse dependent baseline corrections

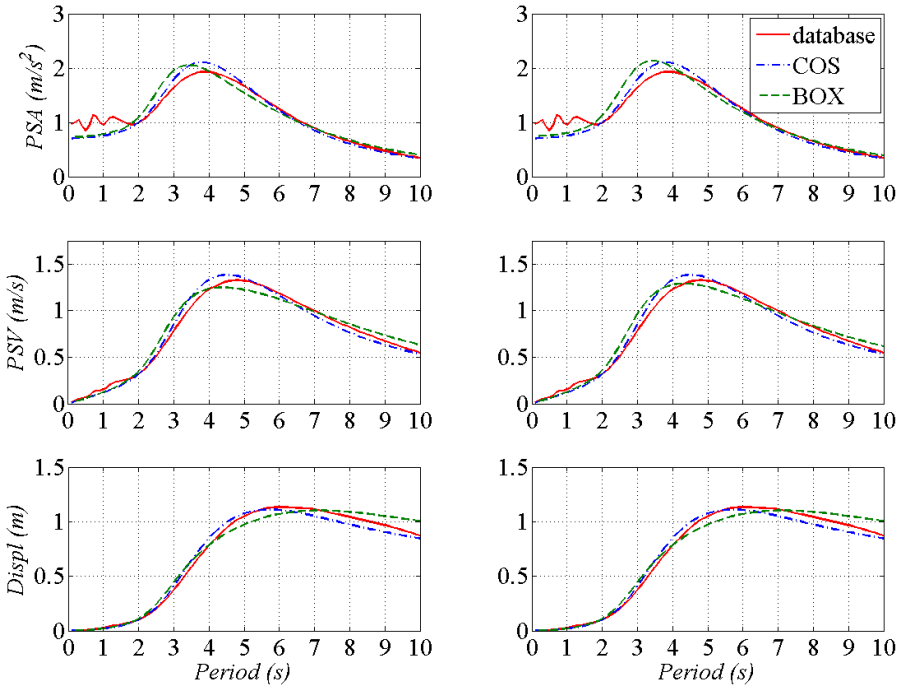

Figure 5.8. Impact baseline corrections on the response spectra – average match for medium pulses (periods between 4s – 6s): (left panel) response spectra when using a uniform baseline corrections; (right panel) response spectra when using a pulse dependent baseline corrections

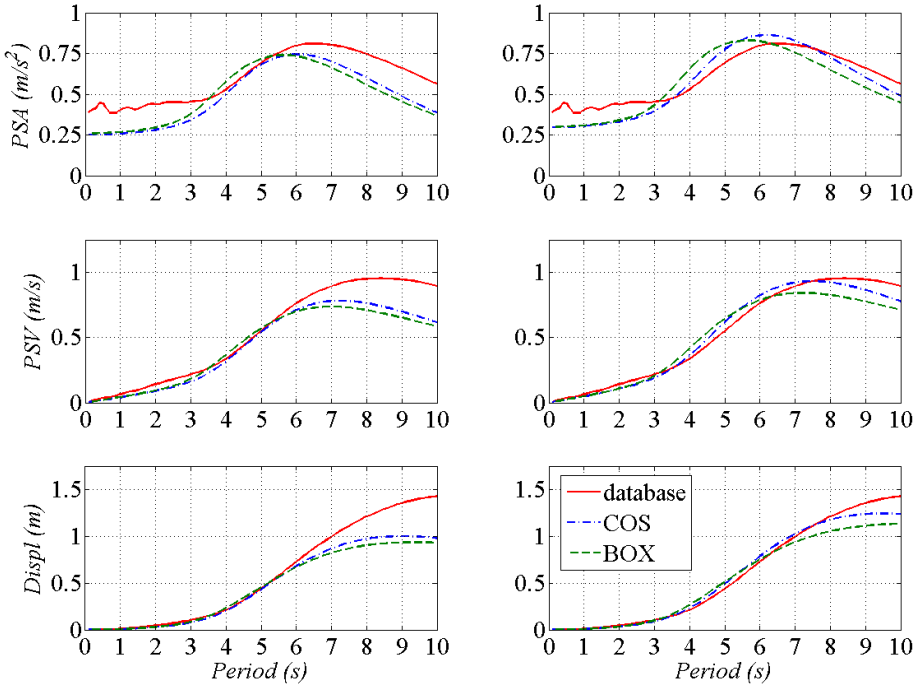

Figure 5.9.Impact baseline corrections on the response spectra – average match for long pulses (periods over 6s): (left panel) response spectra when using a uniform baseline corrections; (right panel) response spectra when using a pulse dependent baseline corrections

5.2.5. Comparison of the simulated and the actual low-frequency content

Following the alternative definition employed for the pulse dominant frequency and the pulse period-dependant baseline corrections employed, the performance of the model is reassessed. The average 5% elastic pseudo-response spectrum across the database is plotted in Figure 5.11. The ratios ρ defined in Eq.(5.13), which were previously employed for the preliminary investigation of the quality of matching across the ensemble, are re-computed. Additionally, the Arias Intensity of the time histories, the cumulative absolute velocity and the PGA, PGV, PGD of the time-histories are also presented, since such indicators for the quality of simulations have been previously employed in the literature (e.g. Conte & Peng, 1997).

Figure 5.10 shows a satisfactory improvement in the values of the ρ ratios in comparison with the values displayed in Figure 5.5. The average response spectra across the ensemble plotted together with the average values across the simulations in

Figure 5.11 confirm these observations. It can be concluded that the proposed pulse model can be calibrated against a given database to satisfactorily simulate on average its structural impact in terms of response spectra.

Further, the potential of the proposed model to generate realistic time-histories is explored in terms of the peak ground accelerations, velocities and displacements. The values in the database are plotted against the mean values of each corresponding pulse process in Figure 5.12. An adequate match is obtained in terms of peak acceleration and velocity values, while in the case of displacements the mean value tends to be overestimated. In Figure 5.13 similar type of scatter plots show the correspondence between the Arias Intensity and the cumulative absolute velocity showing acceptable results.

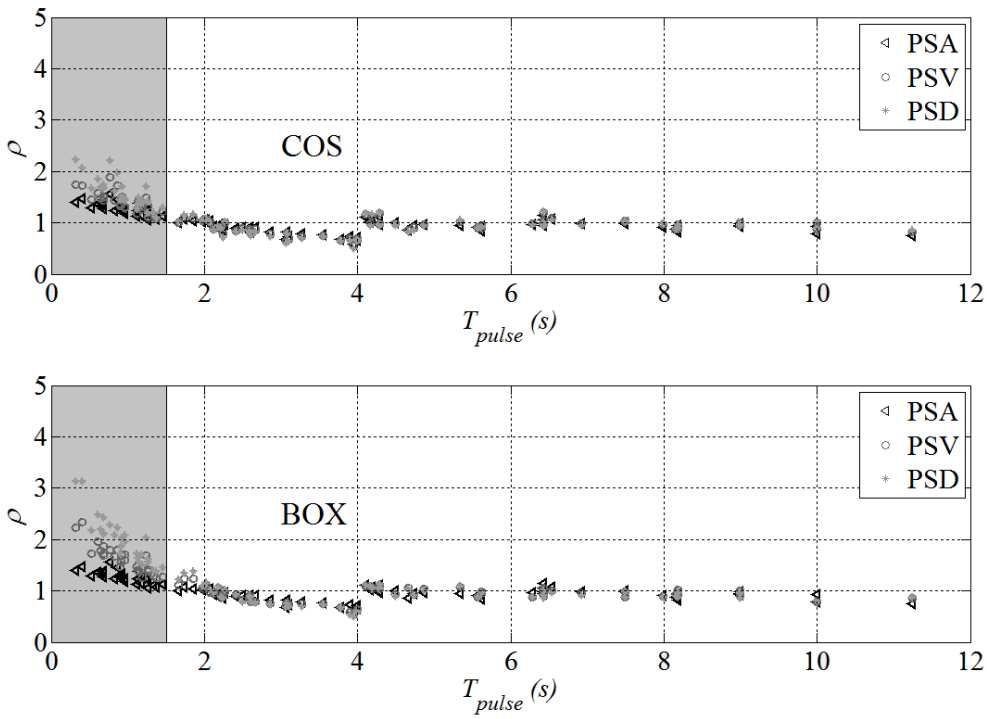

Figure 5.10. Quality of matching across the database after calibration from a response spectrum perspective

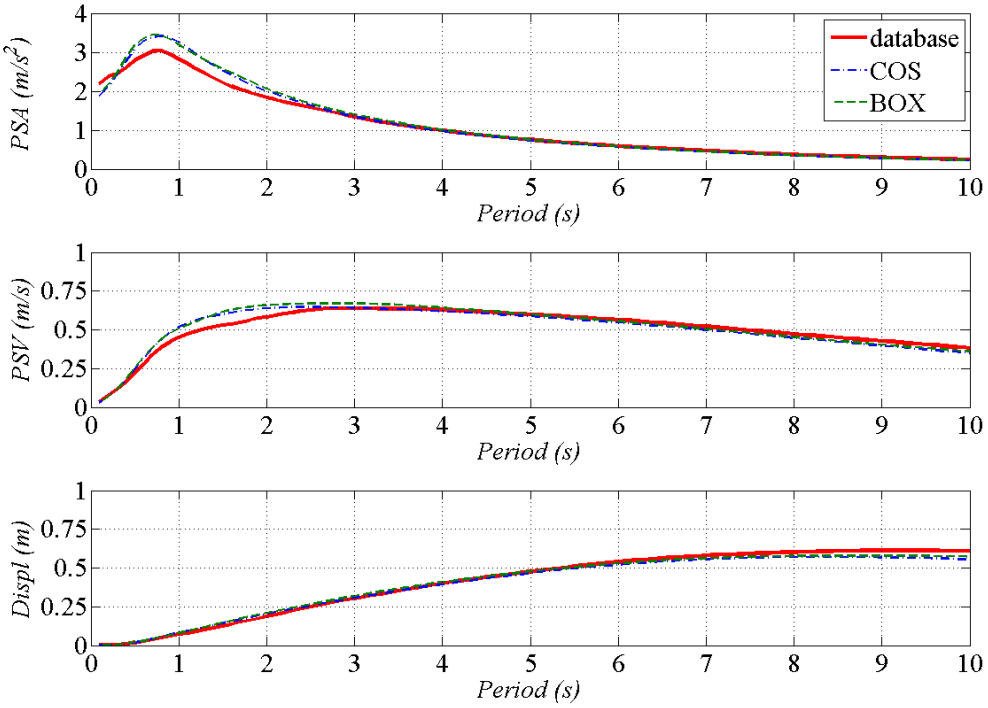

Figure 5.11. Average response spectra across the database considered

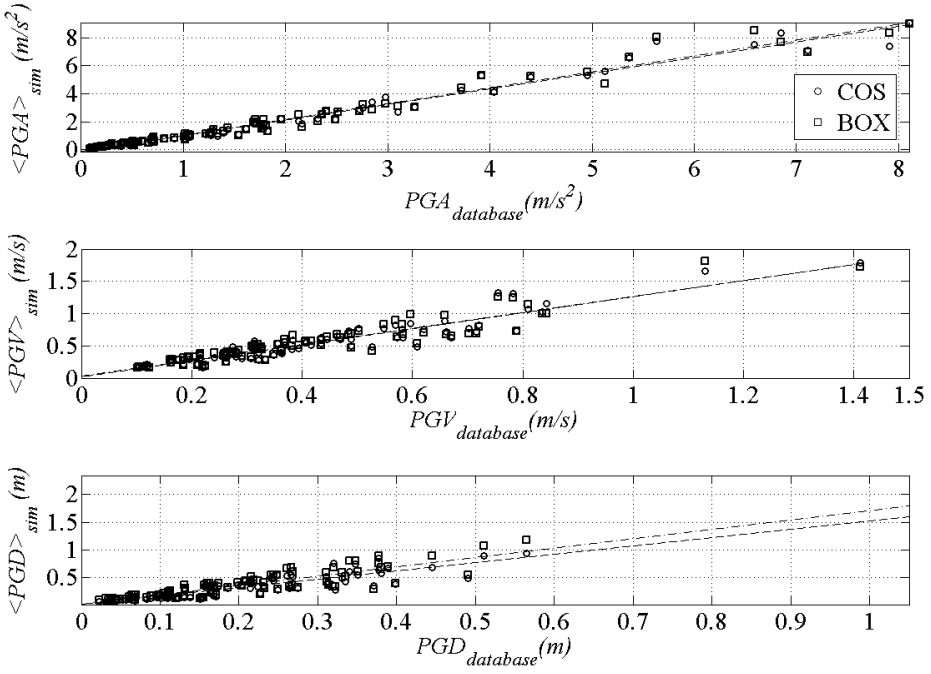

Figure 5.12. Peak ground acceleration, velocity and displacement of the simulations, compared with the values in the database

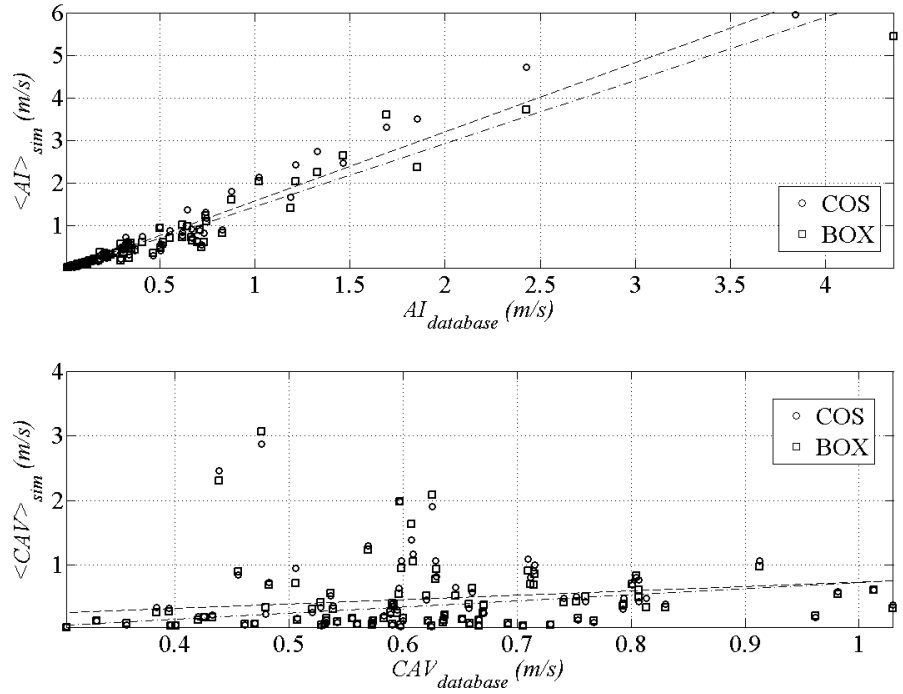

Figure 5.13. Arias Intensity and Cumulative absolute velocity for the simulations, compared with the values in the database

5.3. ACCOUNTING FOR RICH LOW-FREQUENCY CONTENT

The database of pulses considered herein contains one pulse for each record. Accordingly, the simplest case scenario was employed for the representation of the pulses, namely a single uniformly modulated process. While this appears to be a reasonable assumption in general, there are records with richer low-frequency content. For example Mukhopadhyay & Gupta (2013b) provide a list of PLGM records from which two pulses are extracted, rather than one. For such cases more detailed representations are needed in order to obtain accurate simulations.

In Figure 5.14 the response spectra of the 1971 San Fernando record is plotted. It can be seen that the peak elastic displacements are significantly underestimated for periods longer than 4sec when a single low-frequency process is utilised. However, when a more detailed representation of the low-frequency content is sought by considering two low-frequency pulse processes rather than one, an improved representation is obtained.

The proposed low-frequency model can thus adapt to more complex representations of the low-frequency content encountered in PLGM records.

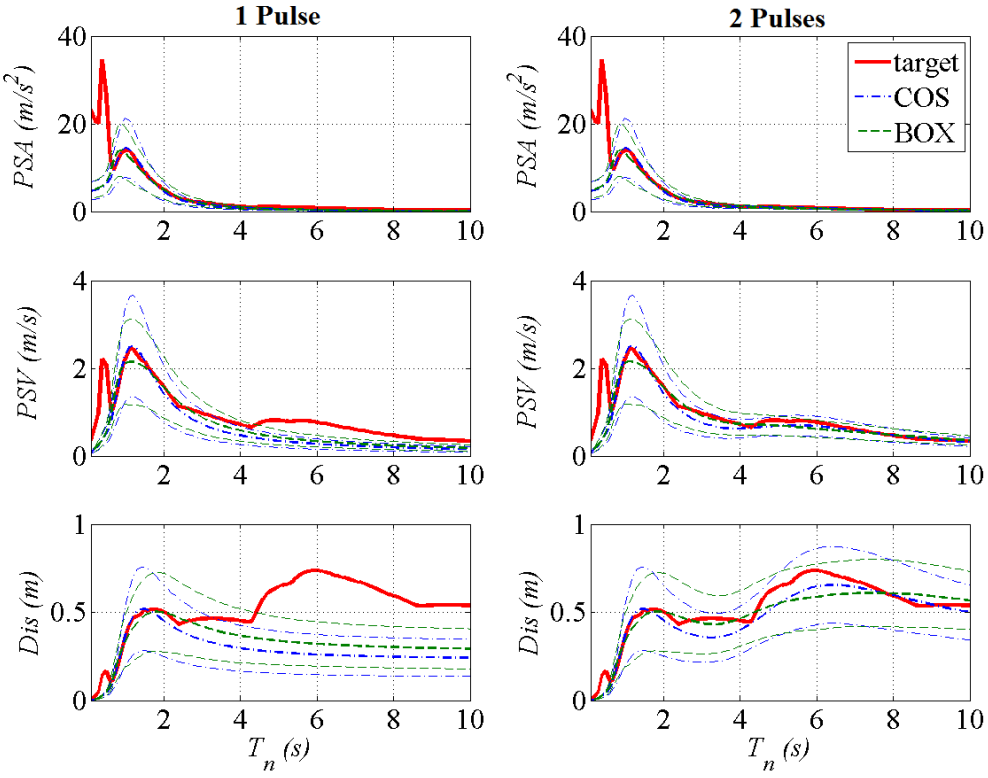

Figure 5.14. Limitations of the database: the one-pulse assumption

5.4. PREDICTIVE EQUATIONS FOR LOW-FREQUENCY MODELS PARAMETERS

Based on the results obtained in the previous sections, recommendations for the definition of the low-frequency model parameters considering arbitrary seismic scenarios in terms of distance and magnitude are provided.

Regression analysis is performed to estimate the relationship between the pulse period and amplitude of the pulse model and the magnitude and closest distance from the rupture for the considered database. Similar empirical relationships for the pulse period and amplitude have been previously derived in the literature for various sets of records (Sommerville 1998, 2003; Alavi & Krawinkler, 2000; Menun & Fu, 2002; Mavroeidis & Papageorgiou, 2003; Bray & Rodriguez-Marek, 2004; Baker, 2007). In Figure 5.15

the relationship between the magnitude of the ground motion and the logarithm of the period is plotted, together with the relationship proposed by Baker (2007). Although the database is the same, a slight difference appears due to the pulse period definition herein considered. The relationship obtained for the pulse period becomes

$$\ln T_p = -6.06 + 1.03 M_W \tag{5.14}$$

Figure 5.16 presents the relationship between the amplitude C_{LF} and the magnitude and distance, given by the formula

$$\ln C_{LF} = 6.82 - 0.004R - 0.995 M_W \tag{5.15}$$

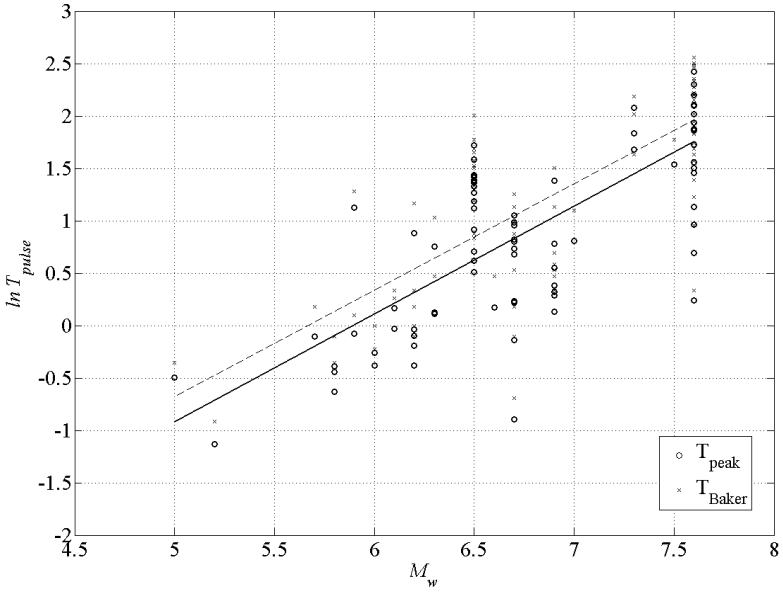

Figure 5.15. Pulse period versus earthquake magnitude

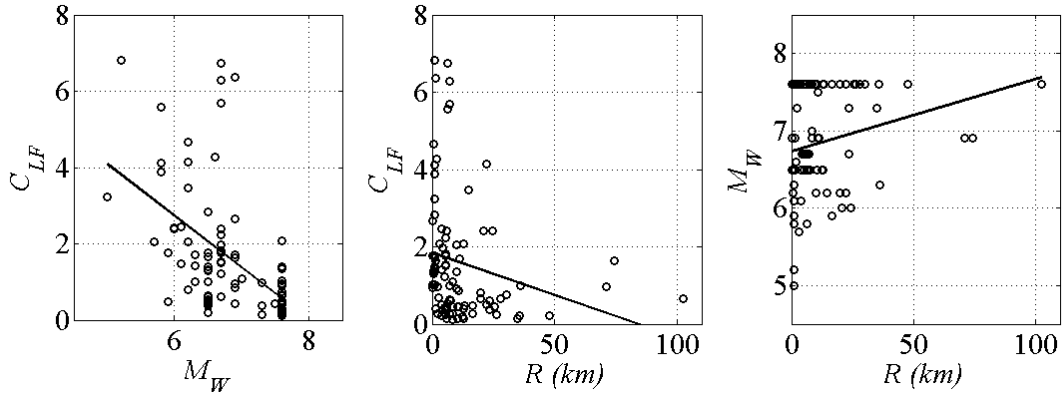

Figure 5.16. Pulse parameters versus magnitude and closest distance

In order to establish what would make a suitable choice for the instant of peak occurrence t_p, the time where 45% of total energy accumulates (Dabaghi et al., 2011) is estimated for the pulses in the database. The values of t_p obtained from fitting the envelope function $a_{LF}(t)$ are compared with the values obtained this way. The scatter plot in Figure 5.17 confirms that the definition remains valid for the case of pulses. The following restriction however needs to be placed when choosing t_p in order to ensure that the resulting time-histories have physical significance, i.e. zero initial acceleration:

$$t_p \geq \frac{T_p}{2} \tag{5.16}$$

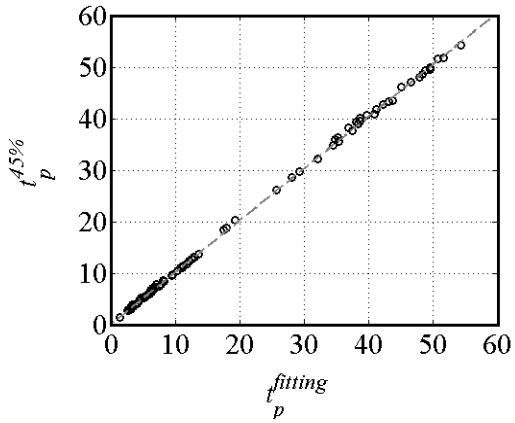

Figure 5.17. The instant of the peak occurrence t_p

Regarding the parameter γ_{LF}, the values obtained after performing the fitting across the ensemble are plotted in Figure 5.18. The values obtained for the considered database

oscillate closely around the average value of 3.4, with a standard deviation around the mean of 0.27. For the definition of the bandwidths B/α and for the appropriate baseline corrections to be applied to the simulated time-histories, the approach presented in the previous section can be employed.

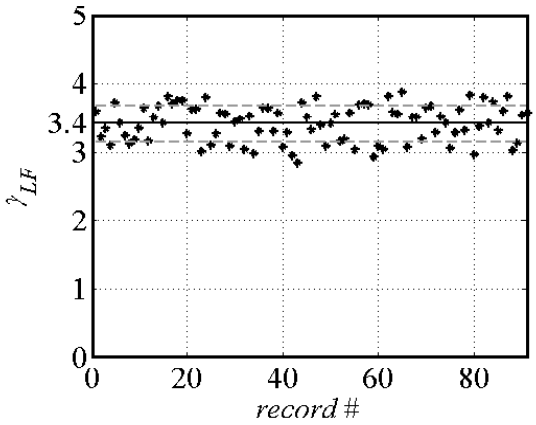

Figure 5.18. The shape parameter γ_{LF}

5.5. USE OF THE PULSE-LIKE GROUND MOTION MODEL TO REPRESENT FIELD RECORDED ACCELEROGRAMS

A field recorded pulse-like accelerogram, belonging to the 1999 Imperial Valley earthquake - station El Centro #6 is considered herein for illustrative purposes of the proposed model potential. The acceleration, velocity and displacement traces are plotted in Figure 5.19. The pulse part and the higher frequency part (residual) of the accelerogram, as separated by Baker (2007) and depicted in Figure 5.19, are treated separately for the calibration of the model.

The Clough-Penzien spectral form is considered for modelling the power spectrum density of the higher frequency (HF) content $G_{HF}(\omega)$. Fitting of the analytic spectrum by using nonlinear least squares regression is performed in order to obtain the values of the parameters. The initial guesses for the parameters ω_g and ζ_g characterizing the soil conditions are estimated using the empirical formulations provided by Lai (1982). The

two alternative LF power spectrum shapes proposed herein are considered for modelling the pulse.

The parameters for the definition of the envelopes are estimated by fitting the analytic functions to the envelopes approximated using the approach presented in the previous section. When performing the fitting for $a_{HF}(t)$, the parameter b_{HF} was fixed at the value estimated based on the effective duration of the residual (Giaralis & Spanos, 2012). For $a_{LF}(t)$ the fixed parameter was the pulse frequency ω_p. The resulting values for the parameters are summarized in Table 5.1.

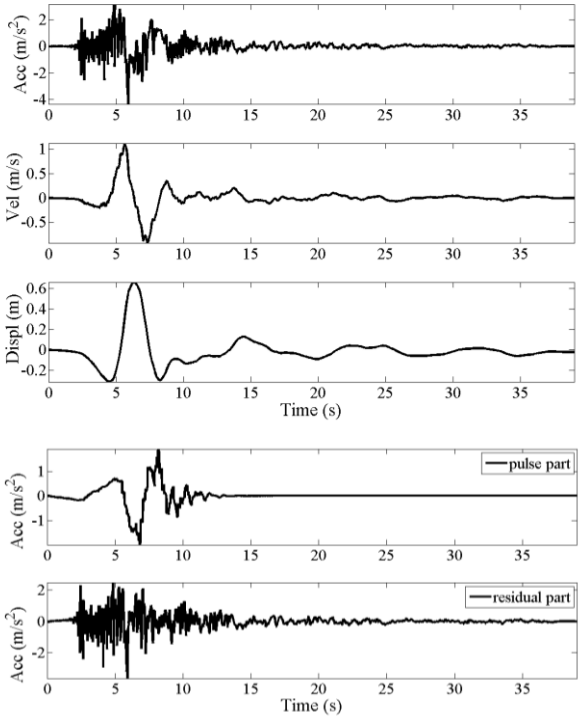

Figure 5.19 Acceleration, velocity and displacement time traces for the Imperial Valley record - station El Centro #6

Table 5.1. Parameters for the simulation of the Imperial Valley – station El Centro #6 accelerogram

High frequency	a_{HF}	$C_{HF} = 0.53$ m/s^2 $b = 0.5$ s^{-1}
	G_{HF}	$\zeta_f = 0.55$ $\omega_f = 2.33$ rad/s $\zeta_f = 0.32$ $\omega_f = 21$ rad/s
Low frequency (pulse)	a_{LF}	$C_{LF} = 1.65$ m/s^2 $\omega_p = 1.65$ rad/s $\gamma = 2.89$ $t_0 = 6.96$ s

	G_{LF}	ω_p = 1.65 rad/s B = 1.65 rad α = 0.50

Ensembles of 200 realizations, compatible with each process (COS, BOX and HF), are generated using the spectral representation method (Appendix C). The time histories have a total duration of 40s and a time step 0.005s. They are subjected to baseline corrections performed by forward/backward high-pass filtering the time-histories using a Butterworth filter (Giaralis & Spanos, 2012). Following the methodology presented in Figure 4.10, the HF samples are superposed to the LF samples resulting in pulse-like time-histories. Arbitrary sample accelerograms corresponding to each process (HF+BOX and HF+COS) and the corresponding velocity and displacement traces are shown in Figure 5.20.

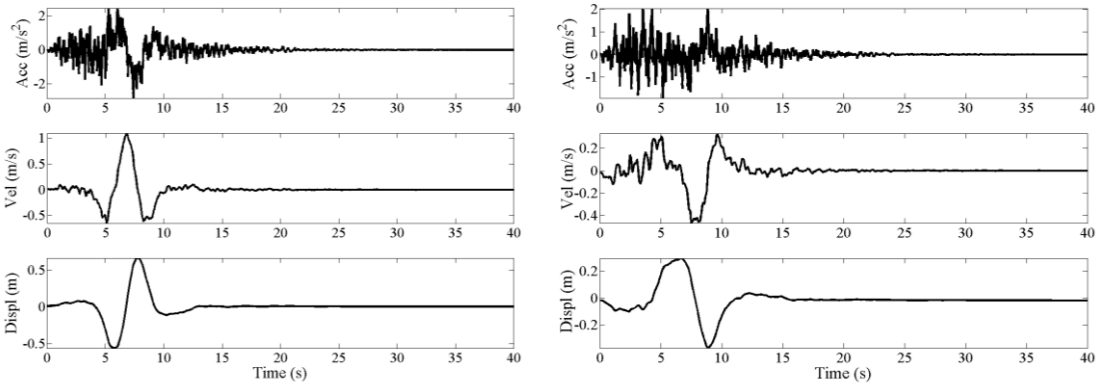

Figure 5.20 Sample of the HF+COS process (left) and of the HF + BOX process (right) generated for the simulation of the Imperial Valley record

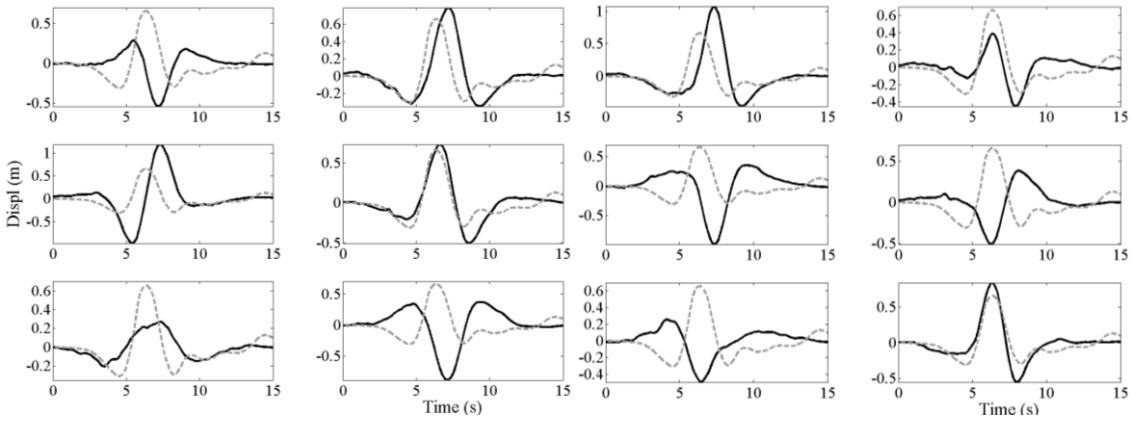

Figure 5.21 Displacement samples of the HF+COS process (left) and HF + BOX process (right) superposed to the Imperial Valley displacement (dashed line)

In Figure 5.21 several synthetic displacement traces are compared with the original time-history. It can be seen that the variability ensured by the model in the characteristics of the pulses i.e. phase and number of oscillations, amplitude, instant of the peak value is taken into account. This qualitative observation shows the applicability of the considered stochastic model to generate realistic PLGM time-histories while ensuring a level of "randomness" within a Monte Carlo analysis context.

Further, the performance of the model in generating structural responses comparable with the recorded accelerogram is investigated. For this purpose, the average elastic and inelastic response spectra for 5% damping ratio corresponding to each process are compared. For the inelastic response spectral ordinates, a ductility factor $\mu = 2$ is considered. This factor represents the ratio between the maximum displacement of a bilinear hysteretic single-degree-of-freedom (SDOF) oscillator and the displacement which causes yielding under a given excitation. The ratio between the initial stiffness and the post-yielding stiffness is taken $\alpha = k_{elastic}/k_{plastic} = 0.05$. Basic statistics (mean and standard deviation) on the spectral responses for each process are superposed to the target values corresponding to the original accelerogram in Figure 5.22. Finally, the following response ratio between the average displacement spectrum of each process and the displacement spectrum of the field recorded accelerogram is defined:

$$\psi(T,\mu) = \frac{E\left[S_d^{sim}(T)\right]}{S_d^{gm}(T)} \tag{5.17}$$

A similar indicator was previously used by Fu & Menun (2004) for assessing the quality of their simulations. The response ratios for the HF+COS process (top panel) and for the HF+BOX process (bottom panel) are shown in Figure 5.23.

A good overall agreement is observed between the spectral responses derived for the simulated time-histories and those corresponding to the target accelerogram. The mean

response spectral values for each ensemble fluctuate closely around the target value regardless of the option LF spectral shape chosen. A better agreement is observed for longer periods (greater than 2sec) in comparison with shorter periods, where the response is controlled by the higher frequency content. An improved matching may be obtained by employing more sophisticated models for representing the higher frequency content which can capture more accurately its characteristics (e.g. Conte & Peng, 1997; Yamamoto & Baker 2011, 2013).

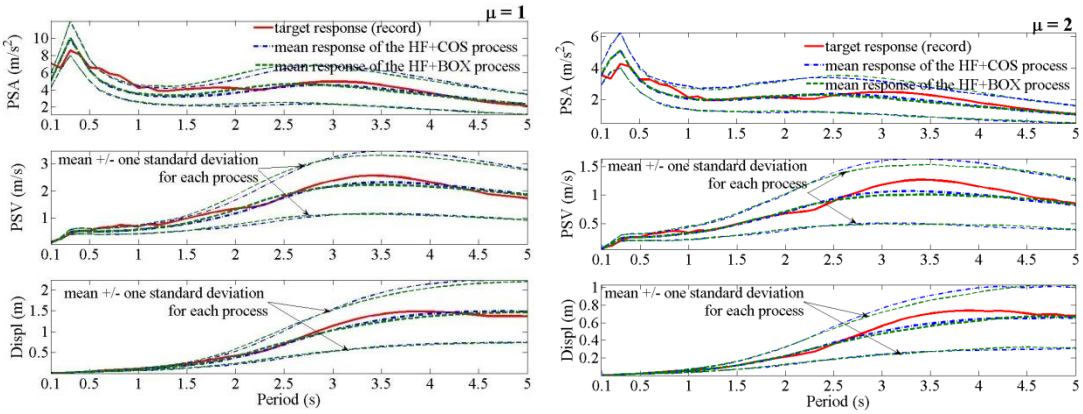

Figure 5.22. Elastic response spectra of the HF+BOX and HF+COS processes (left). Inelastic response spectra of the HF+BOX and HF+COS processes for a ductility factor μ=2 (right)

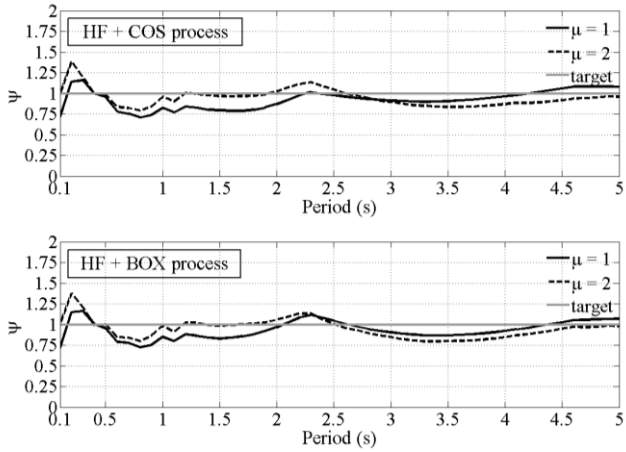

Figure 5.23 Response ratio for the HF+COS and HF+BOX processes

5.6. INCLUSION OF PULSES IN ACCELEROGRAMS COMPATIBLE WITH THE EUROCODE 8 SEISMIC RESPONSE SPECTRA

The extensive amount of damage experiences in the areas affected by PLGMs calls for incorporation of regulations in this regard in the structural design standards. Accordingly, valuable research is devoted to evaluating the probability to experience such ground motions, especially for sites located close to seismic faults (Iervolino & Cornell, 2008; Shahi & Baker, 2011; Chioccarelli & Iervolino, 2013). Further, the incorporation of pulses effects in design ground motions for the purposes of nonlinear response history analysis is also explored (e.g. NEHRP Consultants Joint Venture, 2011; Almufti et al, 2013).

In this context, the herein proposed model is employed for including pulse effects in accelerograms in an alternative manner. Consider a Clough-Penzien evolutionary power spectrum compatible with the elastic response spectrum of the European EC8 aseismic code (CEN 2004) (Giaralis & Spanos, 2012). Assuming this represents the HF content of a pulse-like process, by adding at least one low-frequency process to model the characteristic pulses, a code-compatible pulse-like process is obtained in virtue of Eqn. (4.5). In Table 5.2 the parameters for defining a Eurocode 8 -compatible process for a specific scenario are provided (Giaralis & Spanos, 2012); for generating the pulse part the LF processes derived in the previous section are considered (Table 5.1).

Table 5.2. Parameters for defining the HF content compatible with EC8 spectrum (Giaralis & Spanos 2012)

High frequency **PGA = 0.36g; Soil B; damping ratio 5%**	a_{HF}	C_{HF} **= 0.18 m/s^2** **b = 0.58 s^{-1}**
	G_{HF}	ζ_f **= 0.90** ω_f **= 2.33 rad/s** ζ_f **= 0.78** ω_f **= 10.73 rad/s**

Random time-histories compatible with the pulse-like EC8 and pulse-free EC8 compatible processes are shown in Figure 5.24. By inspecting the peak elastic and inelastic structural responses illustrated in Figure 5.25, the major impact of pulses especially for periods longer than 2s can be clearly observed. It can be concluded that this approach may represent a viable alternative to spectrum matched accelerograms for accounting for pulse effects.

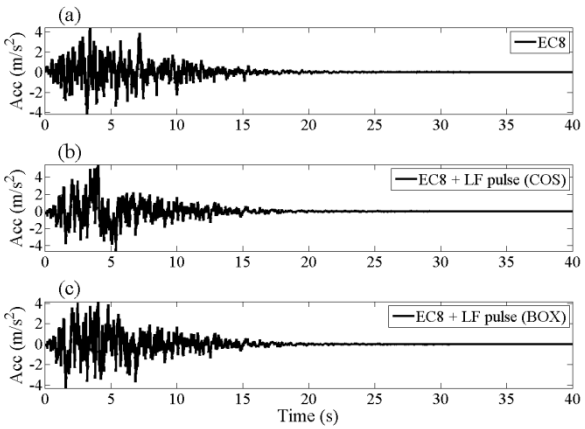

Figure 5.24 Acceleration samples compatible with EC8: pulse-free (a) and pulse-like (b, c)

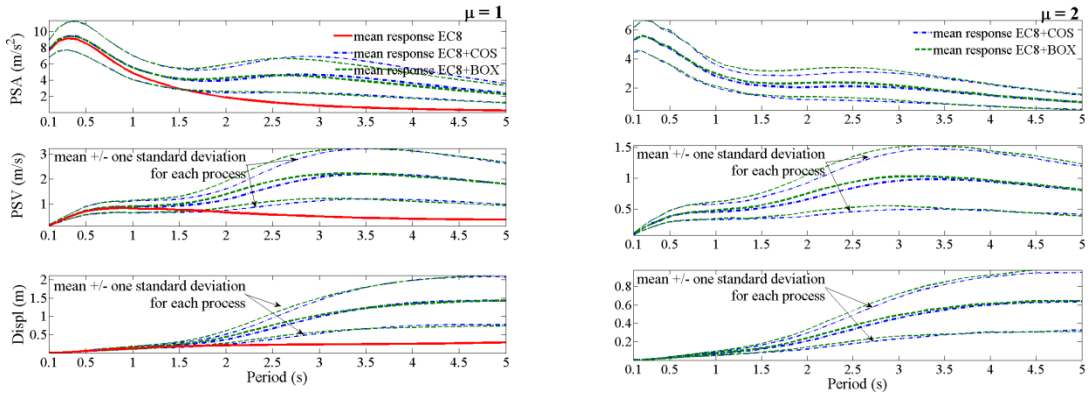

Figure 5.25. Elastic response spectra for pulse-like EC8 compatible processes (left). Inelastic response spectra for pulse-like EC8 compatible for a ductility factor μ=2 (right).

CHAPTER 6 : PERFORMANCE ASSESSMENT OF WAVELET-BASED REPRESENTATION TECHNIQUES FOR THE CHARACTERIZATION OF PULSE-LIKE GROUND MOTIONS

6.1. PRELIMINARY REMARKS

In recent years, wavelet-based and other TFR techniques have been employed for signals' representations on the time-frequency plane in various structural dynamics applications. The review paper of Spanos & Failla (2005) for example provides an overview of the topics investigated using these methods, which range from denoising, characterization or simulation of signals, to damage detection and modal identification. In this context various TFR methods have been also employed for the analysis and characterization PLGM records, as reviewed in Chapter 3. Although TFR techniques are quite mature from the theoretical viewpoint, their effectiveness in signal representation and feature extraction remains case-dependant. Along these lines, comparative assessments of the performance of certain TFRs for analysing recorded PLGMs have been considered in the literature (e.g. Yaghmaei-Sabegh, 2010; Vassiliou & Makris, 2011; Mollaioli & Bosi, 2012). However, the issue of which technique is more appropriate to apply for the analysis of PLGMs remains open. This is caused mainly by some built-in limitations of the performance of TFR techniques when dealing with such signals which, as previously discussed, are characterized by rich and broadband frequency content.

An investigation on the performance of several such methods which have been previously used in seismic related applications, namely the HWT, MWPT, ST and EMD is carried out in this chapter. Contrary to previous comparative studies of limited scope which use specific recorded PLGM records for the purpose, this numerical experimentation uses artificial pulse-like processes/accelerograms, with known properties. Specifically, the PLGM model proposed in Chapter 4 is considered. Two Monte Carlo simulation-based methodologies are considered for assessing the performance of TFR techniques for PLGMs. The first one evaluates the potential of the wavelet-based techniques to characterize the underlying energy distribution of pulse-like ground motion processes, using an approach previously employed in the literature to benchmark TFR techniques for pulse-free ground motions. The second methodology is novel and has been developed for the purpose of this study to investigate the quality of the pulses extracted from synthetic accelerograms by comparing them with the low-frequency content of the PLGM stochastic model. The previously mentioned wavelet-based techniques, together with the EMD, are employed for filtering out the pulses; several indicators are considered for evaluating the impact of the techniques on the quality of the information extracted.

The numerical results obtained reason the use of wavelet-based techniques for characterization of PL accelerograms and, if used with care, for identification and extraction of pulses. This is exemplified by considering three field recorded accelerograms, classified as being of pulse type.

6.2. DESCRIPTION OF THE SYNTHETIC PULSE-LIKE PROCESSES

The performance of the herein considered TFR techniques is assessed by considering artificial time-histories, rather than field recorded, since their main attributes are user

defined and thus known in advance. The pulse model proposed in Chapter 4 is employed for generating pulse-like processes. The regressive relationships developed in Chapter 5 are used for obtaining the values of the parameters considering specific seismic scenarios in terms of magnitude and distance.

Firstly some basic statistics on the pulse periods and magnitudes characterizing records in the herein considered database (Appendix D) are obtained. In Figure 6.1 a probability distribution fitting of the pulse periods in the database is performed. A lognormal distribution offers a reasonable fit for the data considered. Based on this information, the pulses in the database are separated in two groups – the "short" pulses with periods under 4sec which characterize the majority of the PL records (over 50% of the database) and the "long" pulses with periods over 4sec. Accordingly, two scenarios are considered: one representative for the "short" pulses, i.e. an earthquake of magnitude of 6.5 and distance of 10km and a second one, specific to the more extreme cases characterized by longer period pulses, i.e. magnitude of 7.5 at the same distance from the rupture. The parameters for the definition of the pulse-processes are reported in Table 6.1.

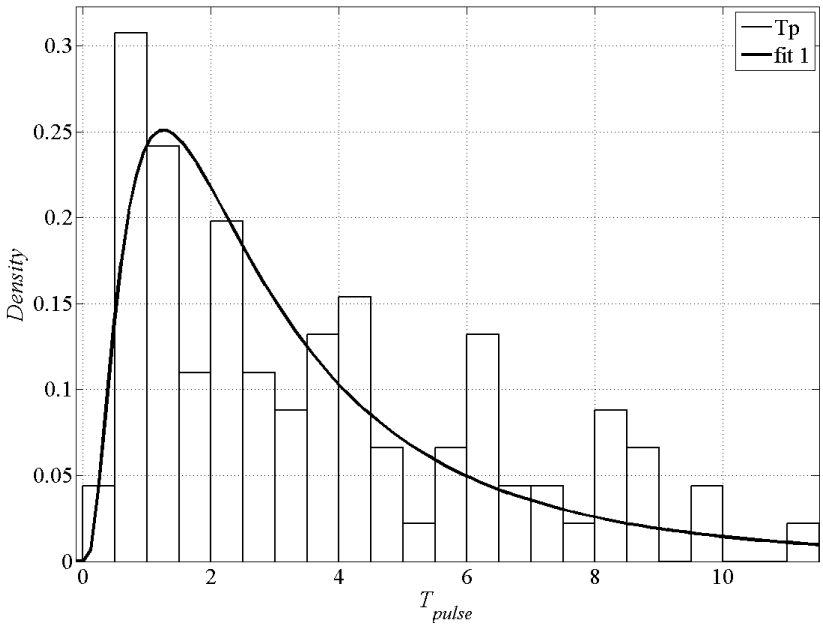

Figure 6.1. Probability distribution fitting of pulse periods across the database considered

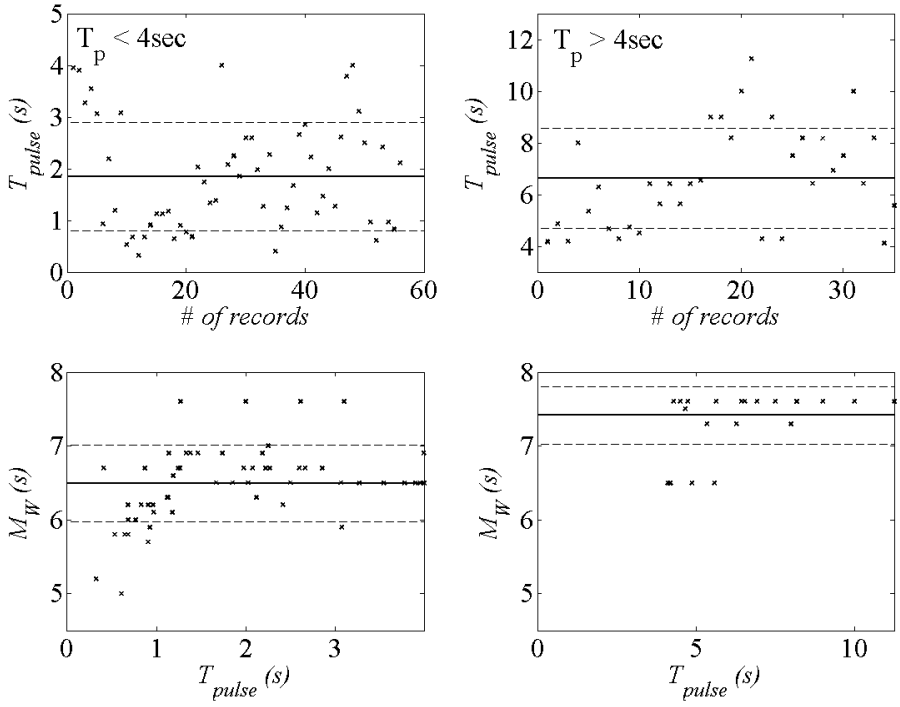

Figure 6.2. Statistics for the pulse periods and magnitudes in the database

For modelling the power spectral density of the pulses, the box-like and the raised-cosine shapes are used. A uniformly modulated Clough-Penzien spectral shape is considered to represent the higher frequency content. The parameters of the high-pass filter ω_f and ζ_f are appropriately chosen to ensure the LF and the HF contents do not overlap. After superposing the LF content with the corresponding HF content as given in Eq. (4.5) and illustrated in Figure 4.5, four pulse-like processes are obtained: two for the short pulses (SCOS and SBOX) and two for the long pulses (LCOS and LBOX).

Collections of 250 pulse-like samples compatible with each process are generated using the spectral representation method (Appendix C). The time-histories have a total duration of 40s and are sampled at a time step *dt = 0.01s*, which allows for a maximum frequency of 314 rad/s. The samples are baseline corrected by forward/backward filtering them with a Butterworth filter of order 2 and a cut-off frequency of 0.2Hz for the SEPSDs and 0.15Hz for the LEPSDs. Following the methodology presented in Figure 4.10, pulse-like accelerograms are obtained by superposing the LF and HF samples.

Table 6.1. Parameters for defining pulse processes

SEPSD M_w = 6.5 R = 10 km	High frequency	a_{HF}	C_{HF} = 0.18 m/s^2 b = 0.58 s^{-1}
		G_{HF}	ζ_f = 0.90 ω_f = 15 rad/s ζ_g = 0.78 ω_g = 10.73 rad/s
	Low frequency (pulse)	a_{LF}	C_{LF} = 2.02 m/s^2 ω_p = 3.32 rad/s γ = 3.40 t_0 = 5 s
		G_{LF}	ω_p = 3.32 rad/s B = 3.32 rad α = 0.50
LEPSD M_w = 7.5 R = 10 km	High frequency	a_{HF}	C_{HF} = 0.18 m/s^2 b = 0.58 s^{-1}
		G_{HF}	ζ_f = 0.90 ω_f = 5 rad/s ζ_g = 0.78 ω_g = 10.73 rad/s
	Low frequency (pulse)	a_{LF}	C_{LF} = 0.7 m/s^2 ω_p = 1.20 rad/s γ = 3.40 t_0 = 5 s
		G_{LF}	ω_p = 1.20 rad/s B = 1.20 rad α = 0.50

6.3. ASSESSMENT IN TERMS OF ENERGY DISTRIBUTION

6.3.1. Methodology

Considering a TFR technique which *locally* and *globally* conserves the energy of the signal (Eq. (2.23)), the EPSD characterizing non-stationary processes can be approximated from the coefficients of the transform in the following way (Spanos & Failla 2004, 2005; Spanos et al., 2005; Liang et al., 2007; Huang & Chen, 2009; Giaralis & Spanos, 2009; Spanos & Kougioumtzoglou, 2012):

$$EPSD(t,\omega) \propto E\left[\left|TFR(t,\omega)\right|^2\right] \tag{6.1}$$

where the operator $E[\cdot]$ represents the mathematical expectation value (i.e. the ensemble average). The degree of accuracy in the previous equation depends on the time localization properties of the analysing functions and on the degree of variation of the envelope. Consequently, it has been shown that the relationship holds for "slowly" varying envelopes (Spanos & Failla 2004, 2005; Spanos et al., 2005).

TFR techniques have been previously employed in the literature for estimating the underlying EPSD characterizing various types of non-stationary processes. Spanos & Failla (2004) compared the performance of three types of wavelets for estimation of analytically defined EPSDs. In a later study of Spanos et al. (2005) harmonic wavelets are assessed in EPSD estimation for artificial and recorded time-histories. Liang et al. (2007) employ the STFT, Morlet WT and EMD in a comparative study for the evaluation of a field recorded accelerogram EPSD. Harmonic wavelets have been used by Spanos & Kougioumtzoglou (2012) for the evaluation of the EPSD in the case of non-linear oscillators' responses.

The MWPT and HWT are employed herein for the estimation of the underlying EPSD of PLGMs (Eq.(4.5)) due to their energy conserving properties (Section 2.4), in order to assess their performance following the methodology presented in Figure 6.3. The average energy distribution of each ensemble of PLGM process realizations is compared with the theoretical EPSD in virtue of the Eq. (6.1). The attention is focused on the potential of each technique to identify the benchmark low-frequency content, in the presence of notable high frequency content as encountered in recorded accelerograms. As discussed in Section 4.2, from a signal processing perspective the latter can be interpreted as high frequency coloured noise corrupting the low-frequency (pulse representative) content. The comparisons are made under the assumptions that the time envelopes corresponding to each component process are varying slowly enough

to ensure the equality in Eq. (6.1) and that a sufficiently large number of EPSD realisations are considered.

Additionally, the ST is considered for the characterization of the energy content of the pulse-like processes by evaluating the average TFR of the ensembles of realisations. It emphasized herein that the ST is not an energy conserving TFR due to the type of normalization employed for the analysing function (see also Section 2.4.3), thus it cannot be used for EPSD estimation.

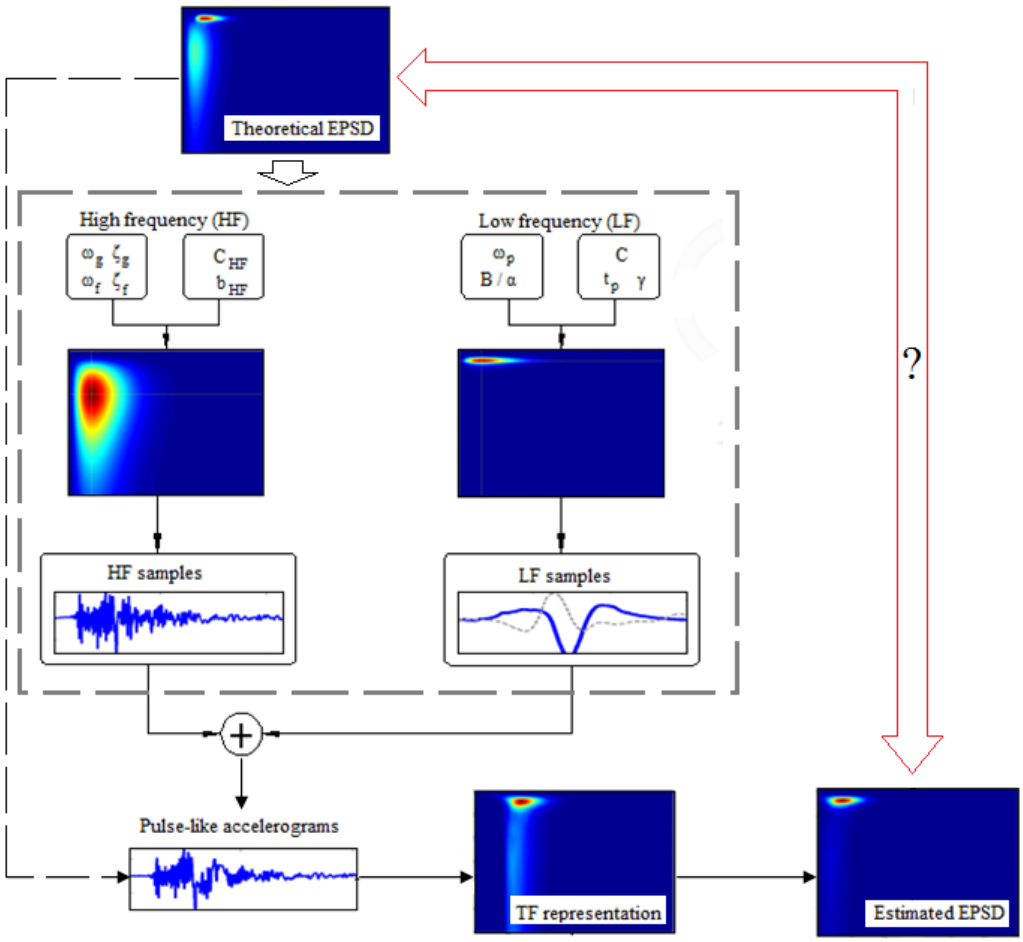

Figure 6.3. Assessment of time-frequency representations via the EPSD

6.3.2. Numerical results

The TFRs of the accelerograms corresponding to each pulse-like process are obtained by means of the ST, the HWT and the MWPT. Prior to being processed with any TFR technique, the samples are zero-padded up to the next power of 2 in order to increase

the speed of the calculations and to diminish the end-effects (Giaralis & Lungu, 2012). In order to obtain satisfactory representations given the broad frequency content of the time-histories, two different resolutions (i.e. levels of detail) are considered when performing the MWPT and the HWT depending on the frequency interval analysed, as shown in Table 6.2. The values for the parameters (window width for the HWT and level of decomposition for the MWPT) are selected upon extensive numerical experimentation considering the quality of approximation in Eq. (6.1) in terms of energy leakage, location of the peaks on the time-frequency plane, locus of the maxima (i.e. ridges) for the LF and HF content and other pertinent criteria.

Table 6.2. Frequency dependant level of resolution employed for each type of wavelet transform

Wavelet Transform	Frequency bandwidth	
	(0 – 10) rad/sec	(10 – 314) rad/sec
HWT	$\delta_\omega = 0.5$ rad/sec ($\delta_\omega = (n_1-m_1)\Delta\omega$)	$\delta_\omega = 2$ rad/sec ($\delta_\omega = (n_2-m_2)\Delta\omega$)
MWPT	Depth of wavelet tree: Level 9 ($\delta_\omega = 0.61$ rad/sec)	Depth of wavelet tree: Level 7 ($\delta_\omega = 2.45$ rad/sec)

Contour plots of the theoretical EPSD, the estimated EPSDs by means of the MWPT and HWT, together with the average TFR obtained using the ST for the four pulse-like processes are pictured in Figure 6.4 - Figure 6.7. Although diverse in appearance, the techniques considered offer satisfactory representation of the energy density distributions in the case of the "short" pulses. Regarding the LEPSDs, characterized by very low frequency energy, the ST offers the clearest representation. An expected side-effect when using wavelet-based techniques is the leakage of the energy content towards higher frequencies, which is sometimes combined with a shift of the predominant

frequencies towards higher values. This is clearly visible for all the representations, being slightly more accentuated in the case of the ST.

In general, the existence of two patches of energy with different locations on the TF plane corresponding to each component process can be identified. This is readily visible for the cases where COS shapes are used for modelling the pulses and for the short processes. In Figure 6.8 and Figure 6.9 the ridges of the HF and LF contents identified by each technique for each process are compared with the target values. For the chosen time-frequency discretizations, the LF content is well localized in frequency. The HF content is closely approximated by the HWT, while the ST presents the most significant shift in peak values.

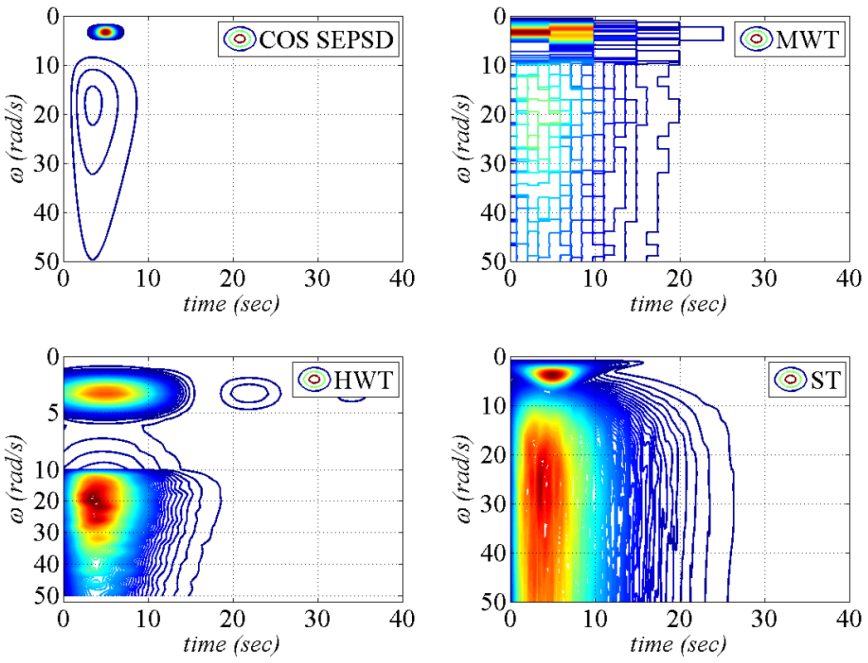

Figure 6.4. Theoretical EPSD for the SCOS process. EPSD estimated via HWT and MWPT. Average TFR using the ST

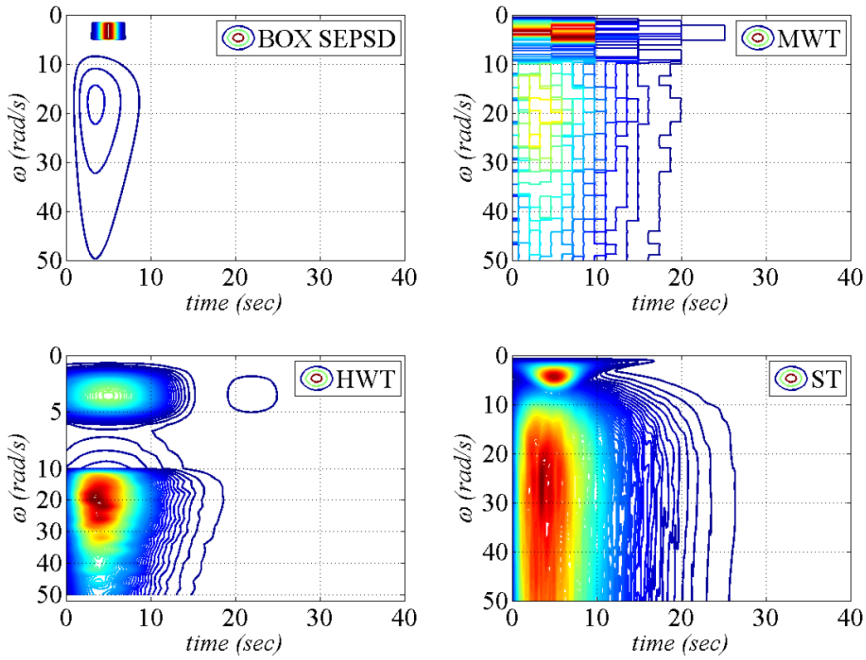

Figure 6.5. Theoretical EPSD for the SBOX process. EPSD estimated via HWT and MWPT. Average TFR using the ST

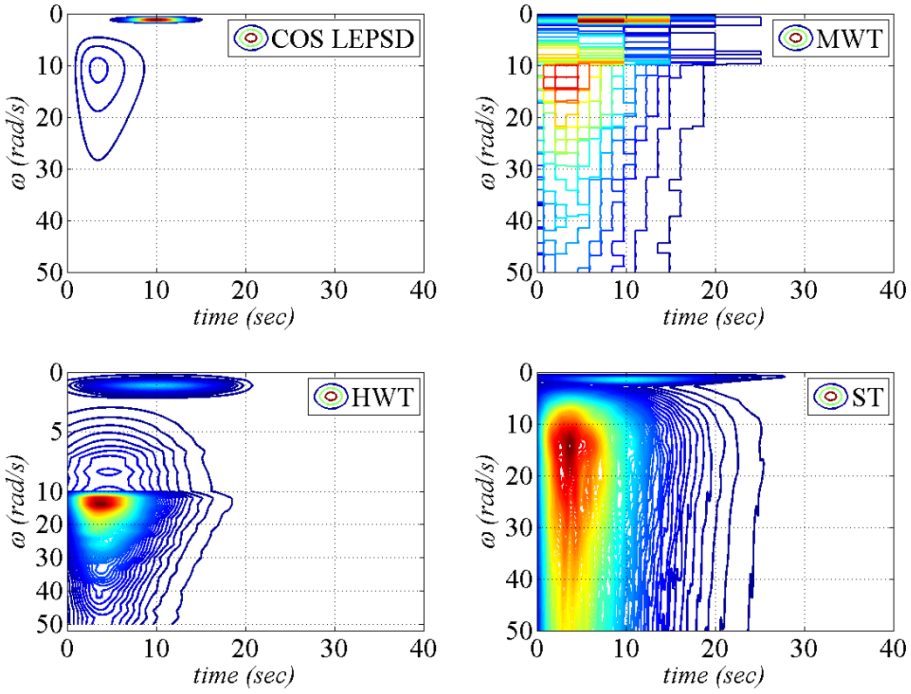

Figure 6.6. Theoretical EPSD for the LCOS process. EPSD estimated via HWT and MWPT. Average TFR using the ST

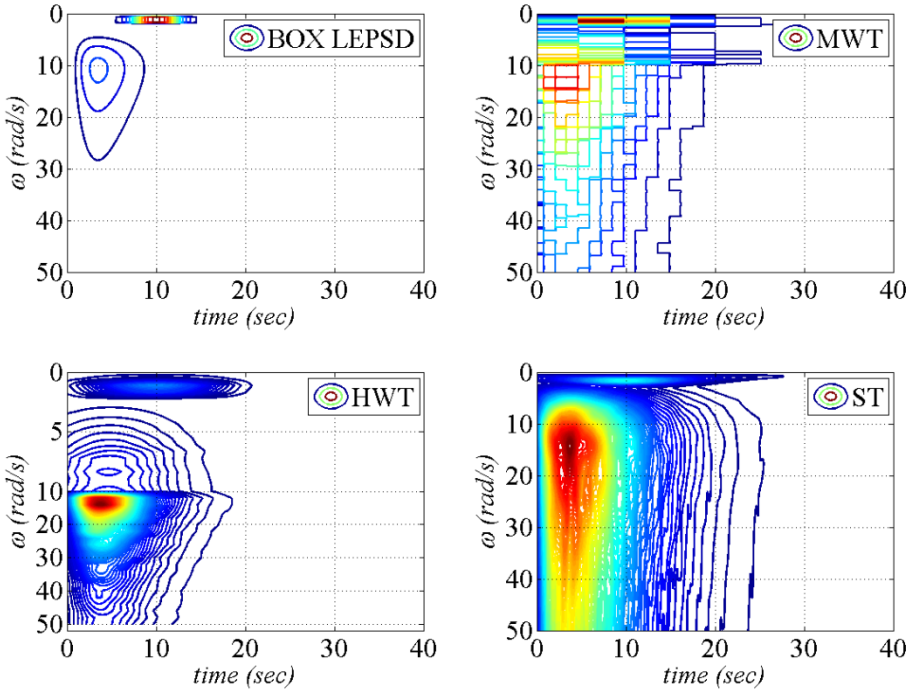

Figure 6.7. Theoretical EPSD for the LBOX process. EPSD estimated via HWT and MWPT. Average TFR using the ST

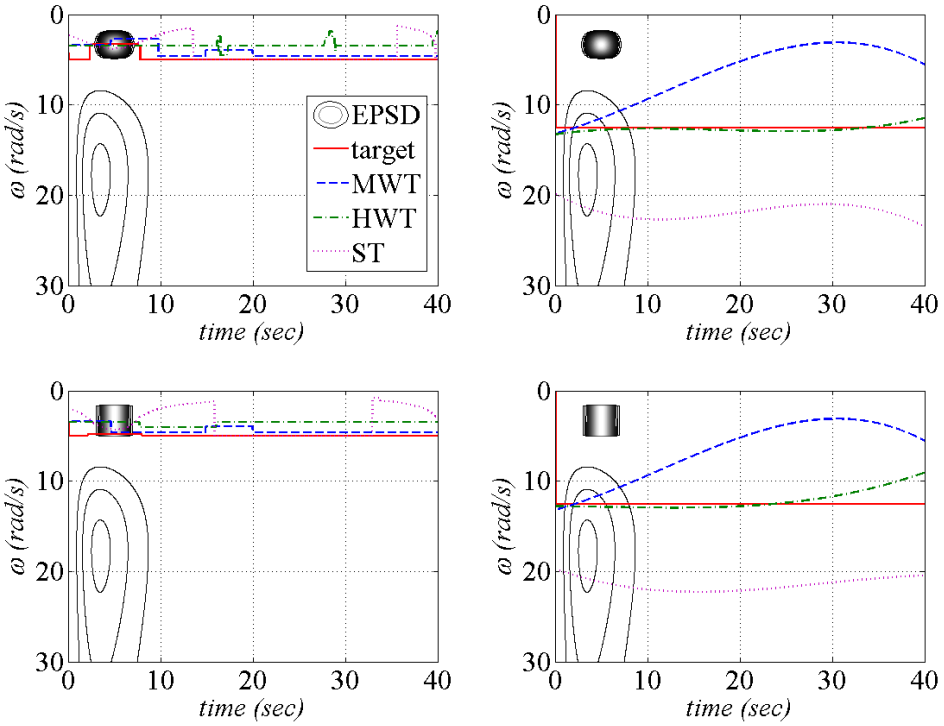

Figure 6.8. Identification of the low-frequency (left panel) and high frequency ridges (right panel) for the short processes: SCOS (top) and SBOX (bottom)

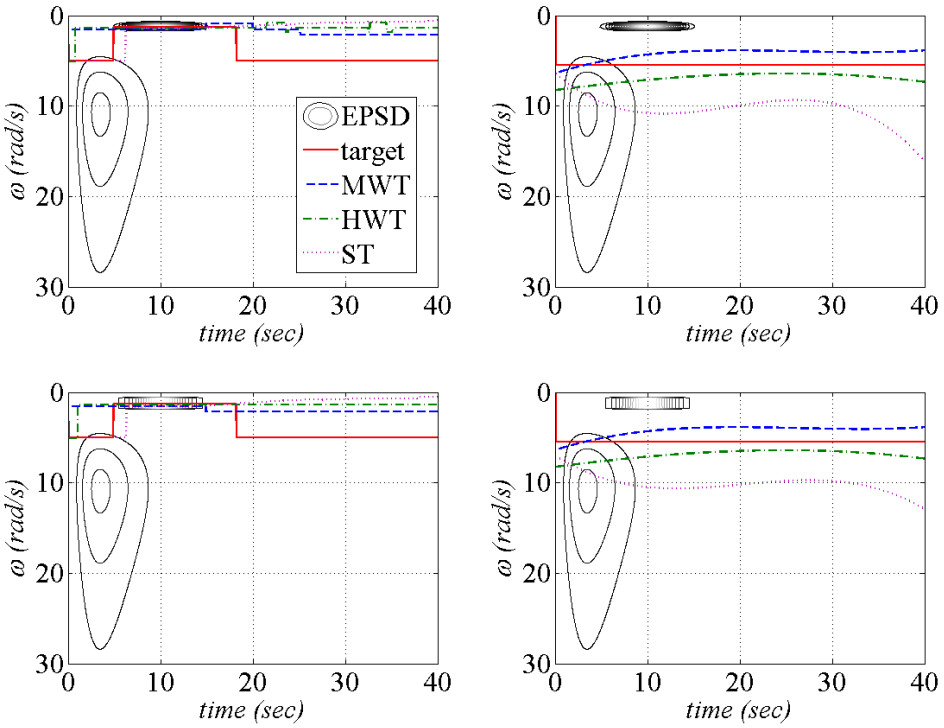

Figure 6.9. Identification of the low-frequency (left panel) and high frequency ridges (right panel) for the long processes: LCOS (top) and LBOX (bottom)

While in general any of the considered techniques performs reasonably well for pulse identification, when very low frequency content exists in the accelerograms the ST achieves better performance in identifying simultaneously the LF and the HF content, while employing a uniform resolution over the entire time-frequency plane.

The GHWT and the MWPT are able to identify the low frequency content in accelerograms if a very fine discretization is used in the corresponding frequency bandwidth. However, for establishing the limits of the bandwidth where such a discretization is needed, the user's subjective input is needed. In order to avoid this, it appears tempting to employ a constant and very fine discretization of the frequency axis; this results into a very broad discretization in time. While in the case of very low-frequencies this is not a significant issue given the short duration of earthquake records, when looking at higher frequencies such long time-windows lead to poor time localization and visibility. This is caused by the fact that the windows used for these

techniques are commonly normalized to have unit energy and thus the high frequency energy is spread over time. The use of a unique resolution is therefore not optimal for broadband signals like the accelerograms; in order to obtain a good visibility over the entire TF plane, several windows sizes need to be used in the analysis. As a matter of fact, this is the reason why pulse identification in generally performed on the velocity time histories, which constitute low-passed filtered versions of the accelerograms (as discussed in Appendix A) and thus have narrower bandwidth (see also Moustafa & Takewaki, 2010).

The results reported herein offer an insight in the potential of the ST, HWT and MWPT for pulse identification in *accelerograms*. Overall, the three techniques have the ability to identify the presence of the very low-frequency characterizing pulse-like processes. Although subjected to limitations, the ST has better performance in identifying simultaneously the LF and the HF content, while employing a uniform resolution over the entire time-frequency plane, which implies minimum subjective intervention. From this perspective, it can thus be used as a preliminary technique for identifying the potential presence of pulses. Based on the information obtained this way, the advantageous bandwidth-dependent detailing offered by the MWPT or the HWT can be utilized for characterization purposes in a reliable way, focusing on the frequency content of interest.

6.4. ASSESSMENT IN TERMS OF RECONSTRUCTED PULSES

6.4.1. Methodology

A second comparison is made to assess the performance of TFR techniques for PLGMs following the novel benchmark methodology detailed in Figure 6.10. As previously presented in Chapter 4 and in Figure 4.10, pulse-like accelerograms are obtained by

superposing the low-frequency samples (LFSs) modelling the "pulses" with the higher frequency samples.

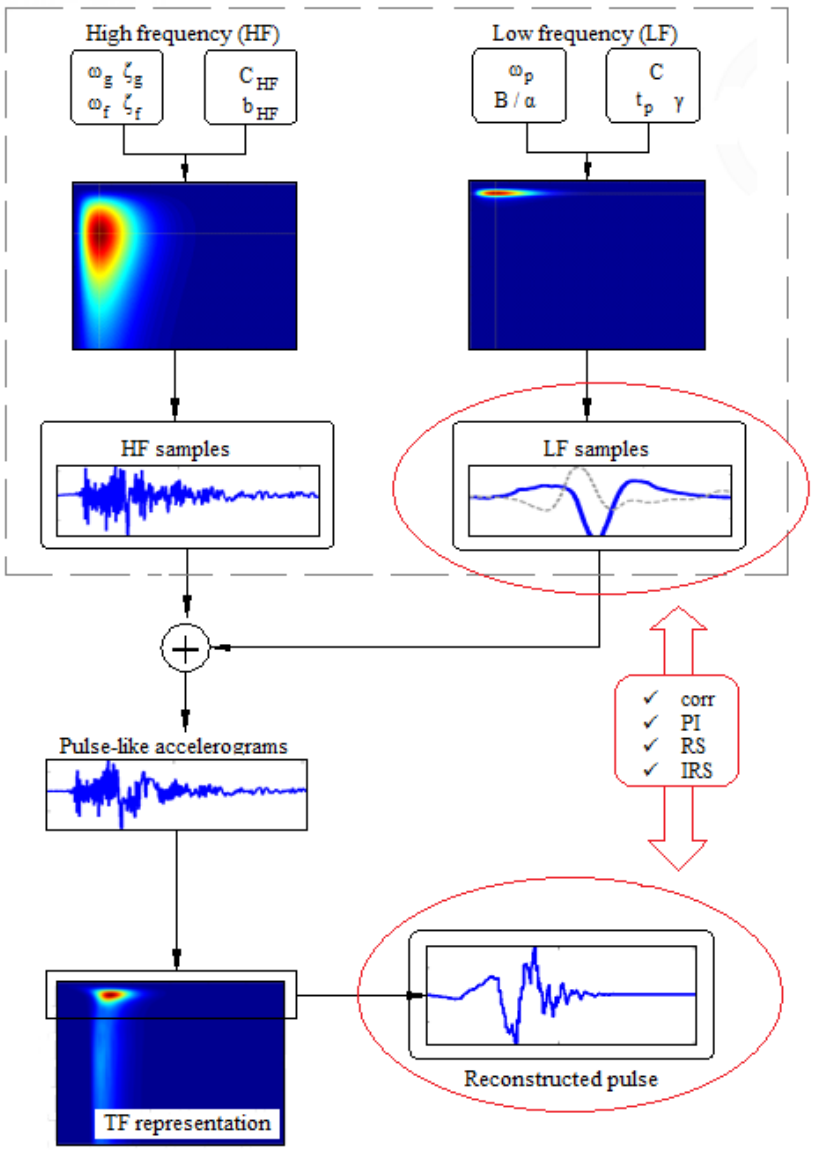

Figure 6.10. Assessment of time-frequency representations via reconstructed pulses (RP)

Herein, the ST, HWT, MWPT and the EMD are employed for extracting the pulses from the resulting accelerograms and their performance is evaluated by comparing these pulses with the original LFSs. The following steps are taken for this purpose:

- The time-frequency decompositions of the analysed signals are performed using each technique.

- The frequency bandwidth expected to characterize to the low-frequency content (i.e. pulses) is established.
- Based on the inversion properties of the considered techniques (discussed in Chapter 2), the signals corresponding to the LF bandwidth are restored. The corresponding signals are referred to as the "reconstructed pulses" (RPs).
- Comparison between the RPs and the LFSs is performed by considering several indicators of the level of similarity. It is expected to obtain this way an insight into the impact of the TFR methods on the synthesized information.

6.4.2. Frequency bandwidth

Based on the statistics obtained for the periods in the database considered in this study (Figure 6.1 and Figure 6.2), in the ensuing numerical experimentations the frequency interval considered to be generally representative for the low-frequency content is set to [0, 5] rad/s. This corresponds to pulse components with periods longer than 1.25 sec.

Although for the artificial processes considered herein the frequency range of the LF process is known prior to the analysis, this type of assumption needs to be made when blindly analysing field recorded accelerograms. The aim herein is to clarify its influence on the accuracy of the extracted pulses and to be taken into account when assessing the performance of the techniques.

6.4.3. Reconstruction of pulses

In the case of the wavelet-based techniques (HWT, MWPT and ST) the RP signals are retrieved from the decompositions using Equations (2.24) and (2.41), by performing the frequency integration over the herein considered low-frequency bandwidth.

Additionally, the EMD is also considered for the reconstruction of pulses since it has emerged in the recent years as a technique suitable for analysing and characterizing

seismic signals. As remarked in previous chapters, the most attractive advantage of the EMD consists in its high adaptability to the data, while one of its most relevant disadvantages appears to be its resistance to statistical analysis. This is mainly caused by the fact that the number of IMFs and their frequency content varies from record to record. Since a large number of records are analysed herein, an algorithm for selecting the IMFs capturing the pulse is needed. After careful numerical experimentation, an energy-based procedure is developed for identifying and isolating the pulses. This choice is motivated by the fact that energy-based approaches have been previously used in the literature to separate the pulse-part from the remaining frequency content (i.e. Loh et al., 2001). The following steps are taken in order to identify the IMFs contributing to the low-frequency content:

1. The total energy of the record and of each decomposing IMF are computed (Eq. (2.3)). The IMFs contributing more than 1% to the total energy of the record are identified:

$$En_{IMF} = \frac{\int_0^{\omega_{max}} \left| \hat{IMF}(\omega) \right|^2 d\omega}{\int_0^{\omega_{max}} \left| \hat{F}(\omega) \right|^2 d\omega} \geq 0.1 \tag{6.2}$$

The amount of energy located in the low-frequency interval (up to 5rad/sec in this case) is estimated:

$$En_{IMF}^{05} = \frac{\int_0^{5} \left| \hat{IMF}(\omega) \right|^2 d\omega}{\int_0^{\omega_{max}} \left| \hat{IMF}(\omega) \right|^2 d\omega} \tag{6.3}$$

The IMFs satisfying the following condition are retained:

$$En_{IMF}^{05} \geq 0.5 \tag{6.4}$$

The RP is obtained as a summation of the selected IMFs. Appropriate baseline corrections are applied to the isolated pulses in order to remove numerical errors induced by the algorithm.

6.4.4. Quality of the reconstructed pulses

6.4.4.1. Degree of correlation

A first indicator used to investigate the accuracy of the information extracted from the accelerograms is the correlation coefficient between the each RP and the corresponding LFS signal. This coefficient indicates the existence of a linear relationship between the considered signals and is given by the equation:

$$Corr\left(RP_i, LFS_i\right) = \frac{E\left[\left(RP_i - \mu_{RP_i}\right)\left(LFS_i - \mu_{LFS_i}\right)\right]}{\sigma_{RP_i}\sigma_{LFS_i}}, \quad i = 1, ..., N \tag{6.5}$$

where N is the number of samples considered, μ stands for the mean, σ for the standard deviation of the samples and $E[\cdot]$ represents the expectation. The correlation coefficient varies in the interval [-1; 1], indicating positive or negative correlation of the analysed data. Perfect linear dependence exists when *Corr* takes the values +/-1.

6.4.4.2. Pulse indicator

Pulse indicators (PI) are parameters which quantify the contribution of the extracted pulses to the pulse-like records. Several such indicators exist in the literature; they are defined based on various properties of the pulses and of the records, like peak ground velocity and energy (Baker, 2007), inner products between the records and the pulses

(Vassiliou & Makris, 2011), intensity of the pulses (Zamora & Riddell, 2011) or energy between consecutive zero-crossings (Mukhopadhyay & Gupta, 2013a). They are commonly employed for classification purposes when analysing sets of records.

Herein, the following PI proposed by Vassiliou & Makris (2011) is used in a different context, namely for TFR assessment purposes. Since no indication of any connection between PIs and the approach used for analysis/pulse extraction procedure is provided by the authors, it is assumed that it represents an appropriate choice.

$$PI = \frac{1}{2}\left(e_a + e_v\right) \geq 0.3$$

$$\textit{where} \qquad (6.6)$$

$$e_a = \frac{\int_{-\infty}^{+\infty} \ddot{u}_{PLGM}(t) \cdot a_P(t)dt}{\int_{-\infty}^{+\infty} \left(\ddot{u}_{PLGM}(t)\right)^2 dt}; \qquad e_v = \frac{\int_{-\infty}^{+\infty} \dot{u}_{PLGM}(t) \cdot v_P(t)dt}{\int_{-\infty}^{+\infty} \left(\dot{u}_{PLGM}(t)\right)^2 dt}$$

In the previous definition the terms $\ddot{u}_{PLGM}$ and $\dot{u}_{PLGM}$ represent the acceleration and velocity traces of the ground motion, while a_p and v_p represent the extracted acceleration pulse and its corresponding velocity trace, which is obtained through integration from a_p. In the ensuing calculations, the PI is first calculated considering the LFSs are the pulses, i.e.

$$a_{P_i} = LFS_i; \qquad v_{P_i} = \int LFS_i dt \qquad (6.7)$$

Secondly, the RPs obtained using each TFR technique are considered for the evaluation of the PI, meaning that

$$a_{P_i} = RP_i; \qquad v_{P_i} = \int RP_i dt \qquad (6.8)$$

The number of time-histories which qualify as being of pulse-type for pulses extracted using the TFR methods is compared with the LFS case, which represents the target case. It is noted herein that similar approaches have been employed by Yaghmaei-Sabegh (2010) and Vassiliou & Makris (2011) for assessing the performance of various wavelets on pulse extraction, by considering field recorded time-histories.

6.4.4.3. Response spectra

In terms of peak structural responses, the elastic and inelastic pseudo-response spectra for 5% critical damping are evaluated and compared for the RPs and the corresponding LFS signals. For the inelastic responses a pre/post yielding stiffening ratio of 0.05 is considered, and the responses for three alternative ductility factors of μ = 2, 4 and 8 are computed. The contribution of the reconstructed pulses RP to the total structural response is evaluated by defining the following pulse response ratio:

$$\psi(T,\mu)=\frac{E\left[S_d^{pulse}(T)\right]}{E\left[S_d^{record}(T)\right]} \tag{6.9}$$

and compared to the target values obtained for the cases of the LFSs.

6.4.5. Numerical results

A total number of 1000 artificial pulse-like accelerograms, coming from 4 different pulse-like processes, are analysed. In Figure 6.11 and Figure 6.12 two such accelerograms, one compatible with the SCOS process (Figure 6.4) and the other with the LBOX process (Figure 6.7) are pictured together with the corresponding velocity and displacement traces. The pulses reconstructed from each accelerogram (RPs) using the TFR techniques are displayed superposed to the original LFS signals. The frequency bandwidth for the reconstruction is set to the interval [0, 5] rad/sec, as previously

discussed. Note that in the case of the MWPT, the actual bandwidth of the pulses is slightly limited by the dyadic discretization of the frequency axis, the resulting pulses having the frequency content [0, 5.49]rad/sec.

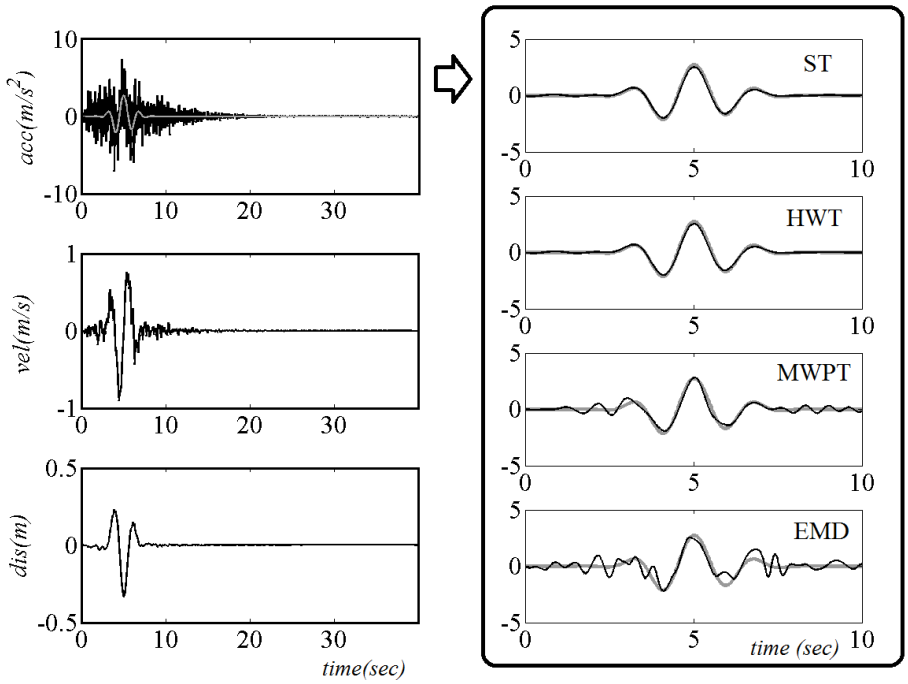

Figure 6.11. Pulse-like accelerogram (arbitrary sample of the SCOS process). Left panel: acceleration, velocity and displacement traces. Right panel: corresponding LF sample (grey) versus reconstructed pulses RPs (black)

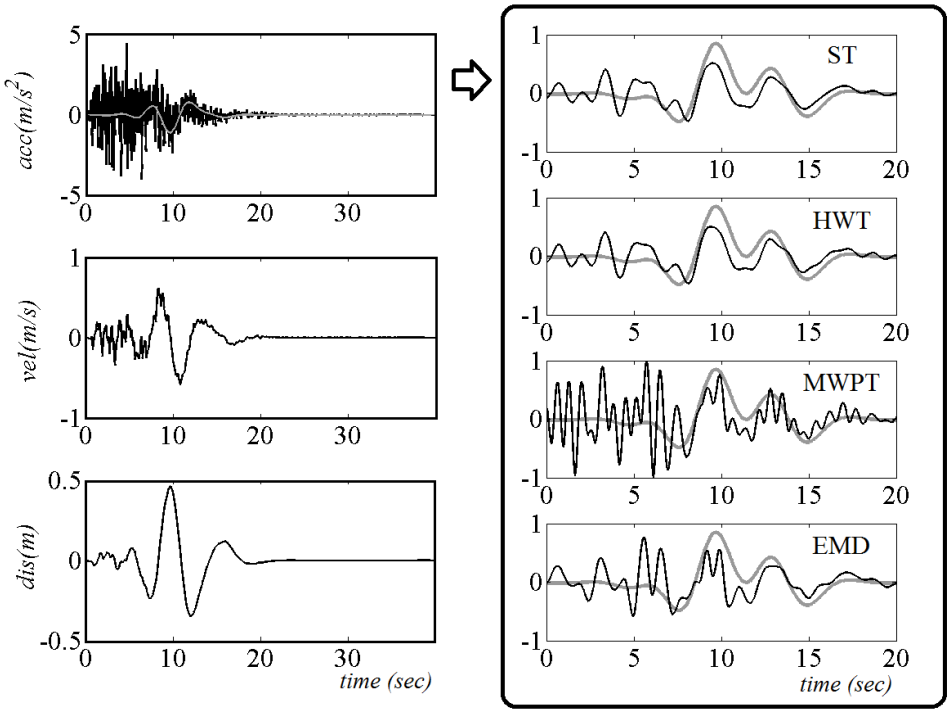

Figure 6.12. Pulse-like accelerogram (arbitrary sample of the LBOX process). Left panel: acceleration, velocity and displacement traces. Right panel: corresponding LF sample (grey) versus reconstructed pulses RPs (black)

The correlation coefficients defined in Eq. (6.5) are determined. Scattered plots of the values obtained for each process and for the case of each technique considered are presented in Figure 6.13÷Figure 6.16. Table 6.3 lists the average correlation coefficient for each case and information on their level of dispersion expressed in terms standard deviation around the mean. High positive correlation values are generally obtained for the "short" pulses, with very good agreement for pulses extracted by means of ST and HWT. This confirms the suitability of the proposed techniques for reconstructing (filtering) pulses from accelerograms and also indicates that the frequency interval considered for this purpose ([0, 5] rad/sec) is relevant for extracting the pulses. Regarding the "long" pulses, the ST and the HWT still offer the closest representation; however the correlation is weaker overall, with higher dispersion around the mean values. This is caused by the fact that for such low frequency content of the pulses, the reconstruction frequency interval of [0, 5] rad/s is too large, allowing for too much higher frequency content compared to the target pulses (see also Figure 6.12). In the case of the MWT, the freedom of controlling the frequency interval is restricted by the dyadic discretization of the time-frequency plane, the pulses extracted are the least similar to those in the input.

Table 6.3 Mean value and standard deviation of the correlation coefficients between the reconstructed pulses and the low-frequency samples

		MWT		HWT		ST		EMD	
		μ	σ	μ	σ	μ	σ	μ	σ
Short	COS	0.90	0.09	0.98	0.02	0.97	0.03	0.88	0.11
	BOX	0.90	0.08	0.95	0.03	0.94	0.04	0.85	0.13
Long	COS	0.48	0.15	0.76	0.13	0.78	0.12	0.58	0.21
	BOX	0.47	0.14	0.75	0.12	0.76	0.11	0.59	0.20

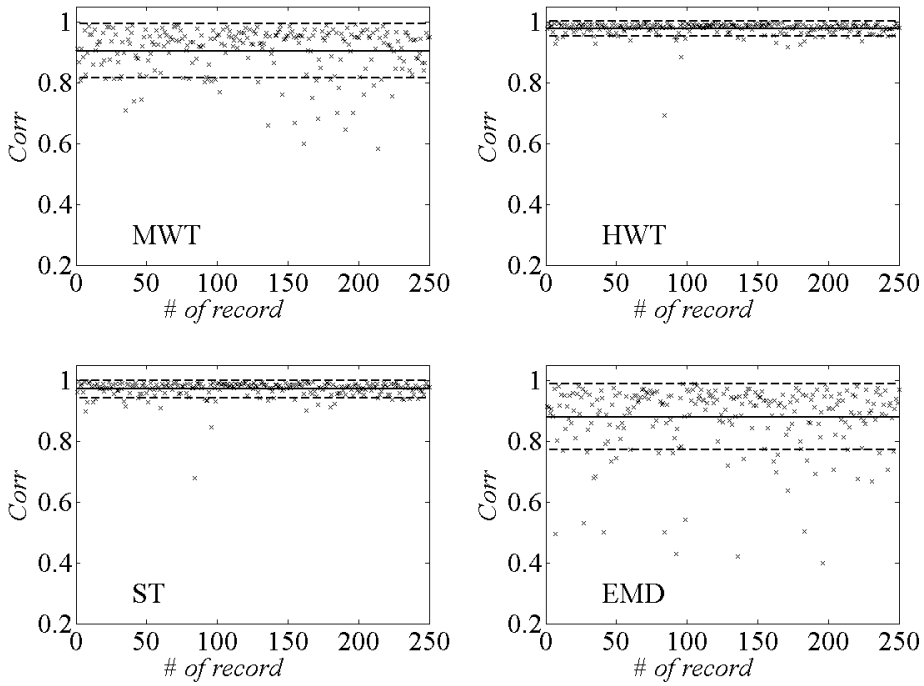

Figure 6.13. SCOS pulses – correlation between the LF samples and the reconstructed pulses

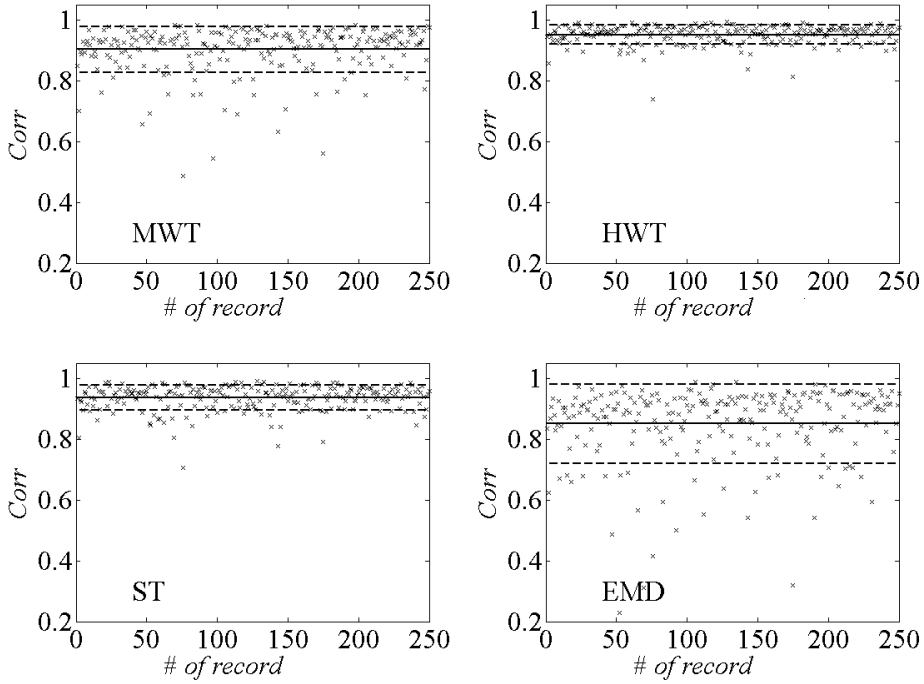

Figure 6.14. SBOX – correlation between the LF samples and the reconstructed pulses

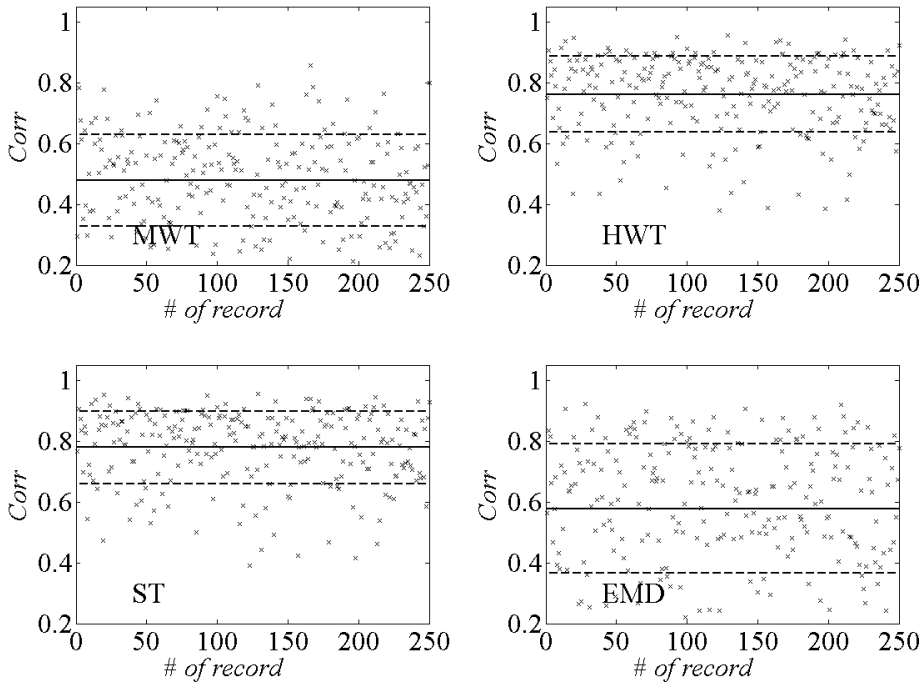

Figure 6.15. LCOS – correlation between the LF samples and the reconstructed pulses

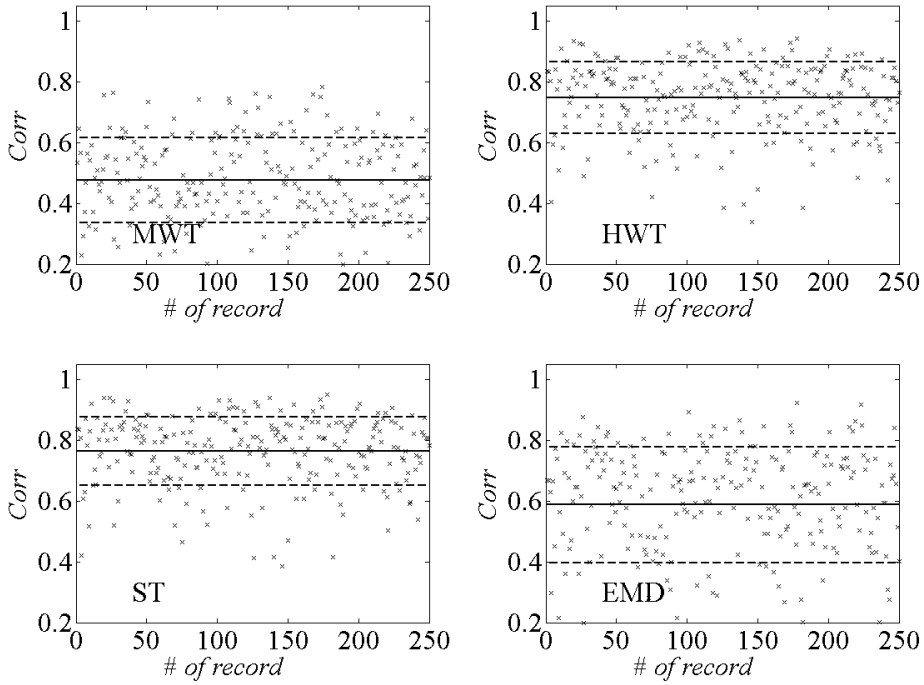

Figure 6.16. LBOX – correlation between the LF samples and the reconstructed pulses

The pulse indicator PI given in Eq. (6.6) is calculated for each time-history analysed, alternatively using the RPs and the LFSs. The percentage of records qualifying as being of pulse-type out of the total number for each case is listed in Table 6.4. The high percentage of records classified as being of pulse-type based on the contribution of the LFS to the accelerograms proves its suitability for the herein considered experiment. The number of records classified as being of pulse-type is very accurate for the case of the ST and HWT and slightly overestimated when using the MWPT for reconstruction.

Finally the contribution of the RPs and of the LFSs to the structural responses is evaluated (Eq. (6.9)). The values of ψ ratios are displayed in Figure 6.17-Figure 6.20 for natural periods ranging from 0.1 to 10 s and for the elastic case ($\mu = 1$) or when various levels of inelastic response are considered ($\mu = 2, 4, 8$).

It can be observed that for periods longer than 2 sec (for "short" pulses) or 4 sec (for "long" pulses) the structural responses are controlled by the low-frequency content and they are well captured by the RPs using the wavelet based techniques. This is expected having in mind that the approach used for extraction is to reconstruct the signal corresponding to the lowest frequency bin. Incongruities appear however around the value of the natural period where the pulse begins to control the structural behaviour. This has to do with the upper limit of the frequency bin considered for the pulse reconstruction, confirming the previous observations regarding its overestimation for the extraction of "long" pulses. As a result, the contribution of the reconstructed pulses to the structural responses is constantly overestimated for these cases. It is thus important to be careful when choosing the interval relevant for the pulse reconstruction. However, once the frequency content of the pulses is established, the corresponding pulse-signal reconstructed using the wavelet based techniques considered herein offers reliable approximation.

In what concerns the RPs obtained using the EMD, the numerical results are relatively poor in comparison with the wavelet-based techniques. This is understandable since it is indeed challenging to establish criteria for the selection of the IMFs which are part of the LF pulse. As an exemplification, while in the case of the artificial time-histories an amount of 50% of the energy located in the [0, 5] rad/s leads to a reasonable approximation of the pulse, when field recorded accelerograms were analysed such a percentage did not ensure an appropriate isolation of the pulse and a smaller energy threshold (i.e. 20%) needed to be set. From this perspective, the numerical experiments performed justify the fact that in the literature the empirical mode decomposition is usually employed for the characterization of specific records, rather than in an automated type of signal processing protocol of the accelerograms.

Table 6.4 Percentage of records identified as being of pulse-type based on the pulses extracted using several time-frequency representation techniques

		LF samples %	MWT %	HWT %	ST %	EMD %
Short	COS	**98**	99.2	98.4	98.4	85.2
	BOX	**98.4**	100	98.8	98.4	83.6
Long	COS	**90.8**	100	96.4	96.0	57.6
	BOX	**96.8**	100	97.6	97.6	64.4

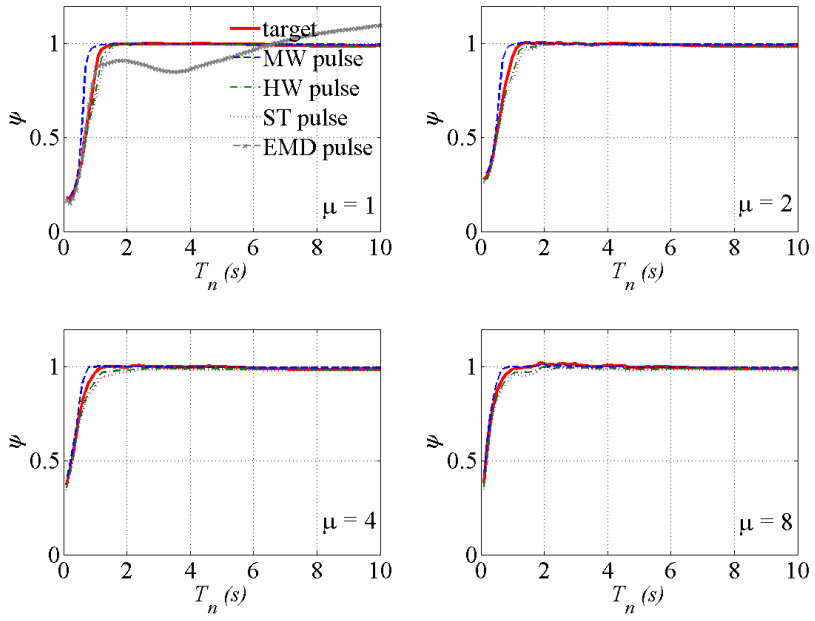

Figure 6.17. Average pulse response ratios for the SCOS samples for ductility factors μ = 1, 2, 4, 8

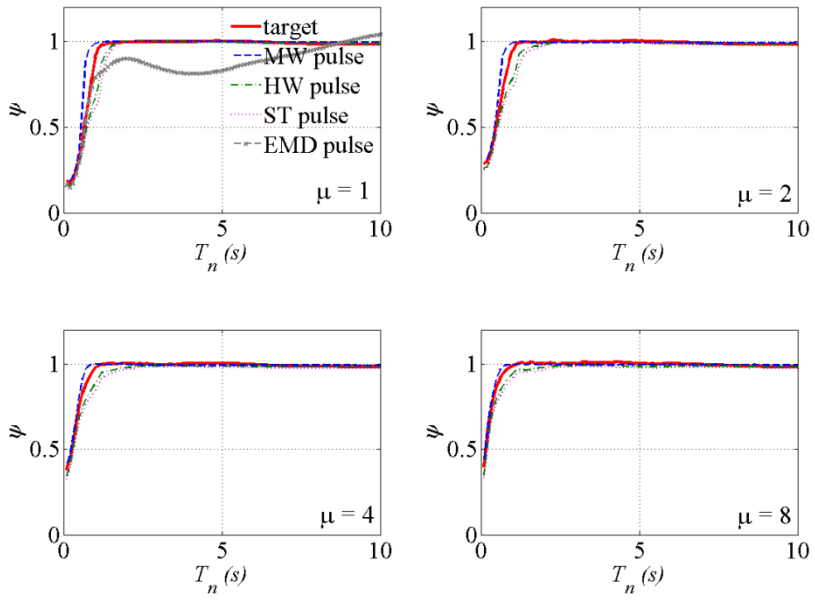

Figure 6.18. Average pulse response ratios for the SBOX samples for ductility factors μ = 1, 2, 4, 8

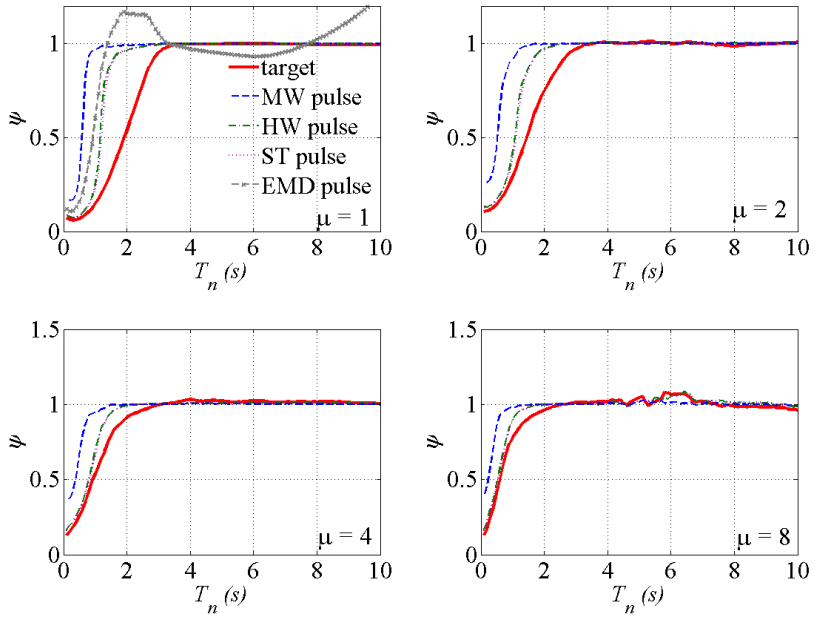

Figure 6.19. Average pulse response ratios for the LCOS samples for ductility factors μ = 1, 2, 4, 8

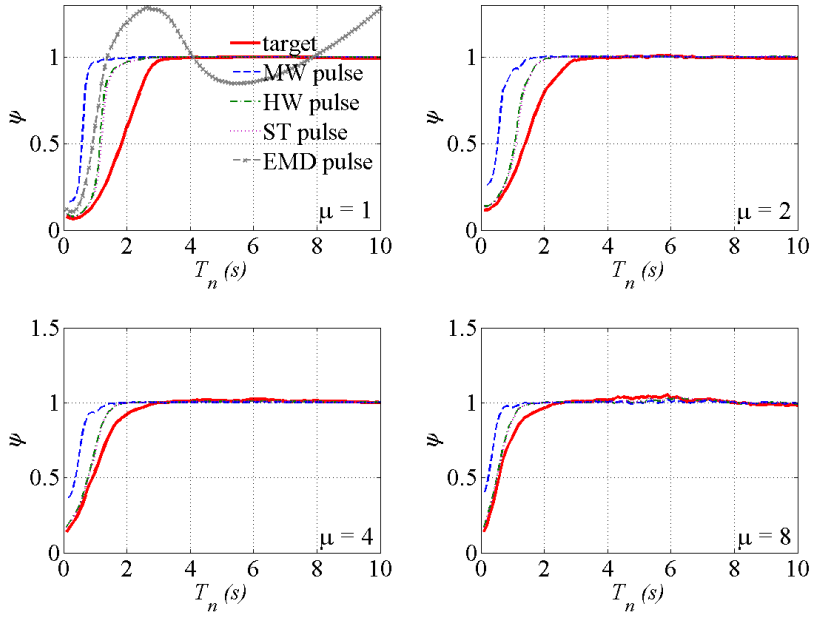

Figure 6.20. Average pulse response ratios for the LBOX samples for ductility factors μ = 1, 2, 4, 8

6.5. FIELD RECORDED ACCELEROGRAMS

The wavelet-based techniques considered herein are employed for extracting the low-frequency content from three field recorded pulse-like accelerograms: the 1979 San Francisco (station Pacoima Dam), the 1989 Loma Prieta (station Saratoga) and the 1994 Northridge (Rinaldi Receiving Station). Figure 6.21, Figure 6.25 and Figure 6.29 display the time-domain and frequency domain representations of the acceleration, velocity and displacement traces; the TFR obtained for the accelerograms using the ST and the energy distributions obtained by means of HWT and MWPT are also shown.

The pulses are extracted from these time-frequency representations of the records by reconstructing the sub-signals corresponding to [0, 5] rad/s bandwidth and can be seen in Figure 6.22, Figure 6.26 and Figure 6.30 together with the corresponding velocity traces. The pulses approximated by Mavroeidis & Papageorgiou (2003) and Baker (2007) for each case, are also plotted for reference.

In the case of the San Fernando and Loma Prieta earthquakes, the velocity traces corresponding to the pulses extracted from the accelerograms offer a very good match to the original velocity traces. In comparison with the fitted pulses, the reconstructed pulses offer a more compete image, accurately capturing all the LF oscillations. This is an expected outcome since they are obtained by low-pass filtering the actual accelerogram.

Further the 5% elastic response spectra (i.e. $\mu = 1$) and the 5% inelastic responses considering the ductility factors of $\mu = 2$, 4 and 8 and a 0.05 elastic-to-plastic stiffening ratio are calculated. In Figure 6.23-Figure 6.24, Figure 6.27-Figure 6.28, Figure 6.31-Figure 6.32 comparative plots for each type of pulse are displayed.

Notice that in the case of the San Francisco and Loma Prieta records, a very good estimation of the structural responses for the cases where the response is controlled by the pulse is obtained. The reconstructed pulses offer a significantly better approximation

that in the case of the pulses obtained fitting wavelets to the velocity traces (Baker, 2007), due to the fact that the entire low frequency content in extracted. A less accurate approximation is obtained for the Northridge record, where the first peak in the spectral responses is not captured by the reconstructed pulses. This implies the frequency interval taken as being representative for the low-frequency content is not broad enough to capture all its effects for the case of this record and should be extended. By looking at the velocity traces in Figure 6.26, the presence of a higher frequency component overriding the low-frequency pulse and not incorporated in the reconstructed pulses is observed.

Regarding the extraction of the pulses from recorded time-histories, the approach commonly used (with some exceptions, e.g. Mukhopadhyay & Gupta, 2013) is to fit various waverforms to the time-histories or, alternatively, their response spectra to the record's pseudo-velocity spectra. The purpose is to use the thus extracted pulses either for the calibration of proposed pulse models or, in other cases, for investigating their impact on structural responses. The identification/extraction is commonly made on the velocity trace. The velocity represents a low-pass filtered version of the accelerogram using a cut-off frequency which depends on the sampling step Δt, thus it is record dependant. In other words, the information in the velocity trace can be retrieved in the accelerogram, by zooming-in in a specific low-frequency interval. This has been proved by reconstructing the signal corresponding to a specific low frequency interval of the accelerogram, comparing it with the defined pulses and inspecting its impact on the elastic and inelastic response spectra. The advantage offered by this approach is that it ensures a constant frequency bandwidth for all the records analysed and thus a uniform analysis of structural responses for various records. From the structural responses point of view, it ensures more realistic and comprehensive evaluations since the information existing in the time-histories is used, rather than approximations which can be limited in

the cases when the low-frequency content is richer than expected, as it is the case of the San Fernando earthquake record for example.

However, as it is concluded from the previous numerical results, establishing a frequency interval or a cut-off frequency generally representative for pulses can be quite challenging. From this perspective an approach for classifying the low-frequency content in PLGMs considering the structural natural frequencies of interest might be more effective.

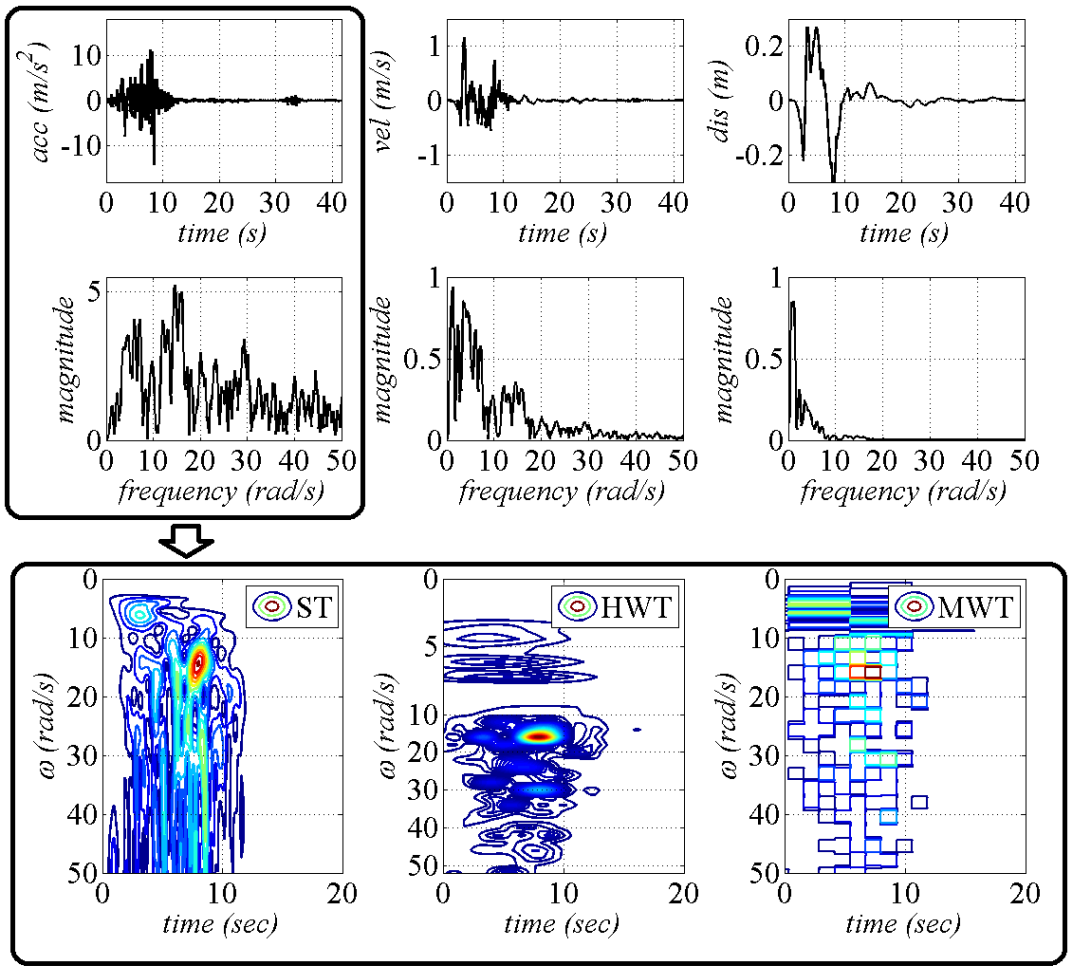

Figure 6.21 1971 San Francisco (Pacoima Dam): acceleration, velocity and displacement traces time domain and frequency domain representations.

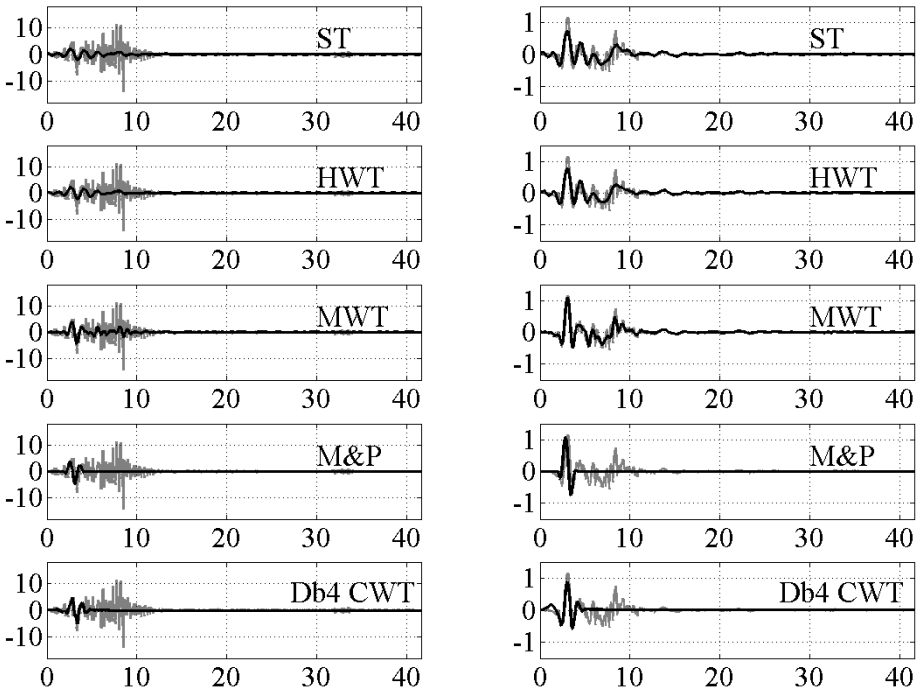

Figure 6.22 1971 San Francisco: reconstructed pulses (acceleration and velocity time-histories)

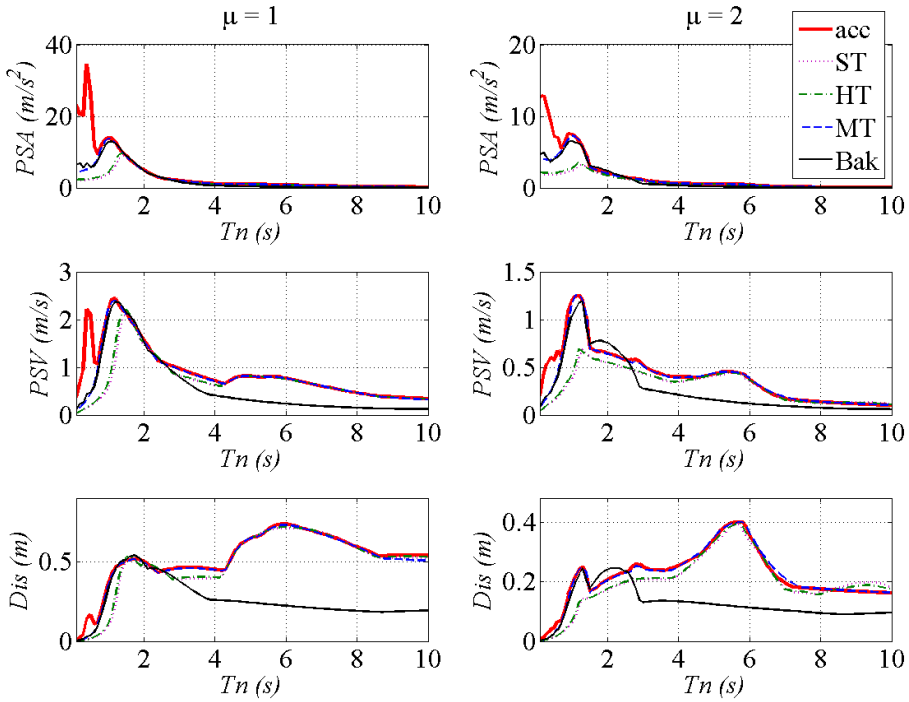

Figure 6.23 1971 San Francisco: spectral responses for constant ductility factor μ = 1 (elastic) and μ = 2

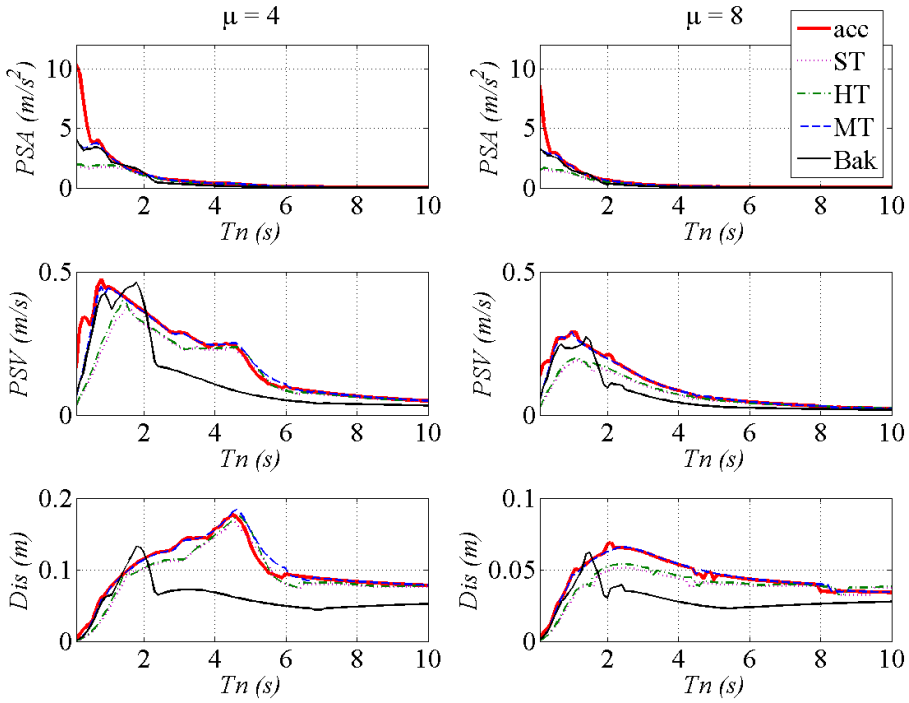

Figure 6.24 1971 San Francisco: spectral responses for constant ductility factor μ = 4 and μ = 8

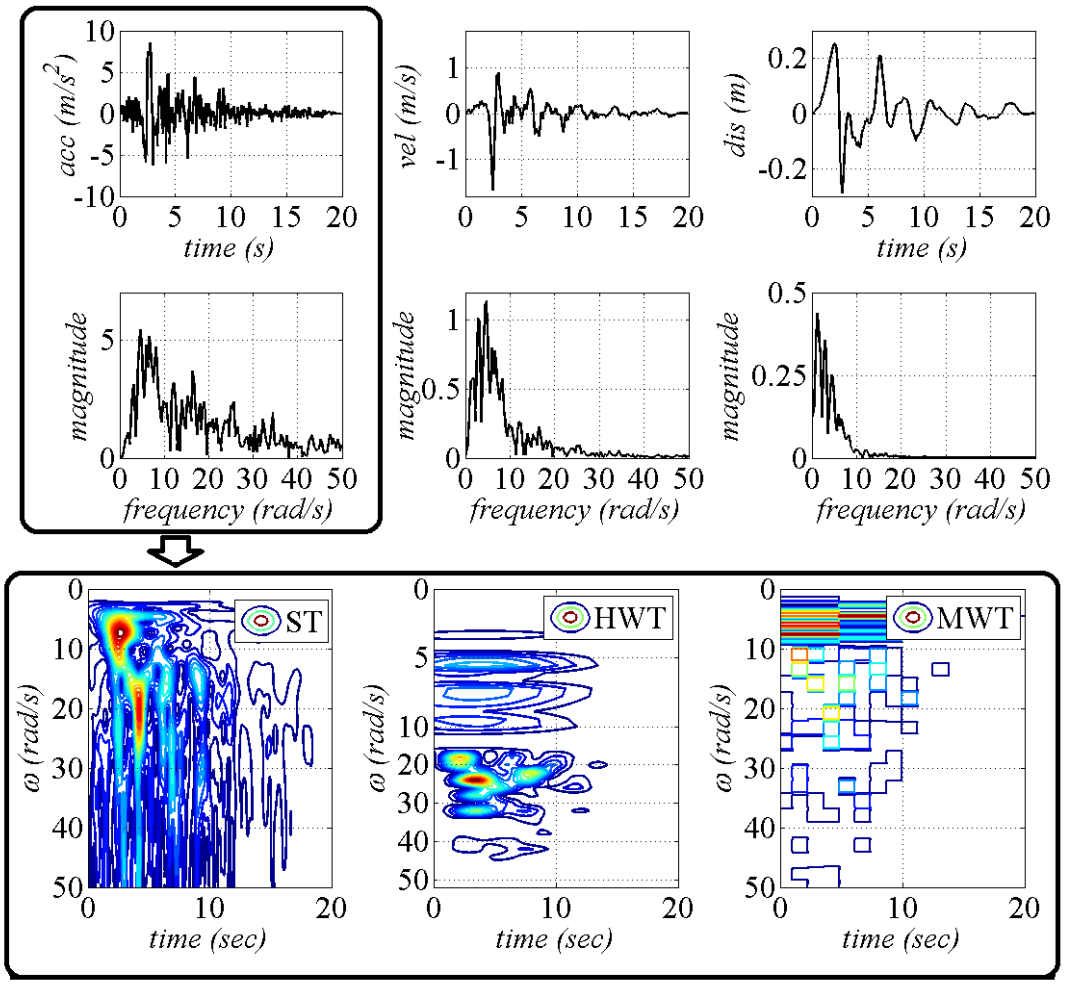

Figure 6.25 1994 Northridge (Rinaldi): acceleration, velocity and displacement traces time domain and frequency domain representations.

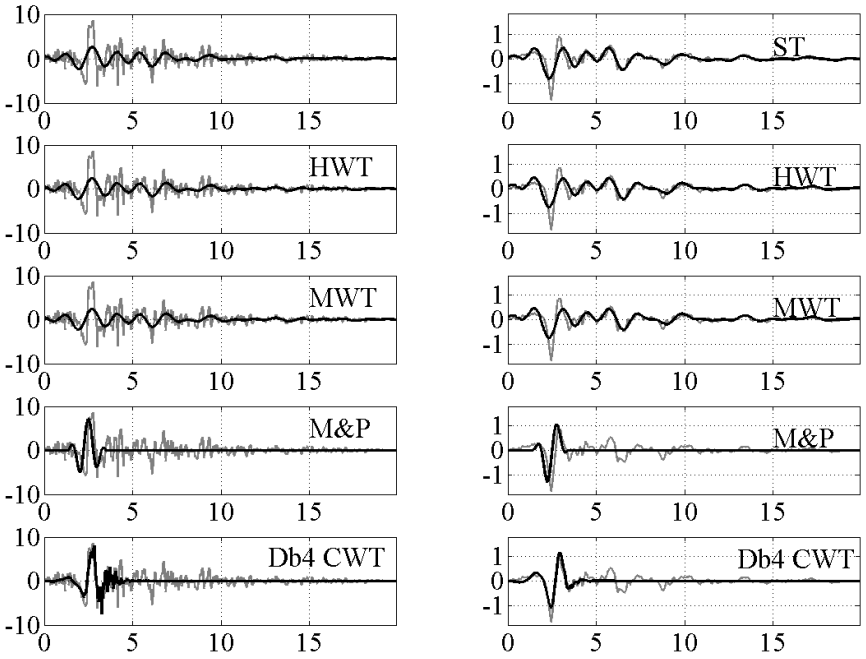

Figure 6.26. 1994 Northridge: reconstructed pulses (acceleration and velocity time-histories)

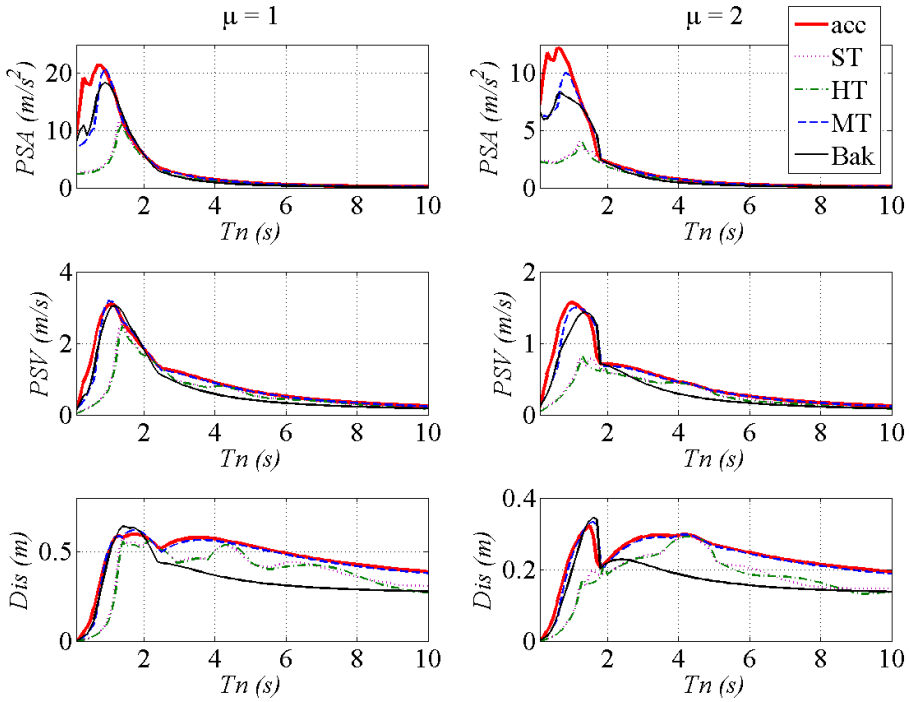

Figure 6.27 1994 Northridge: spectral responses for constant ductility factor μ = 1 (elastic) and μ = 2

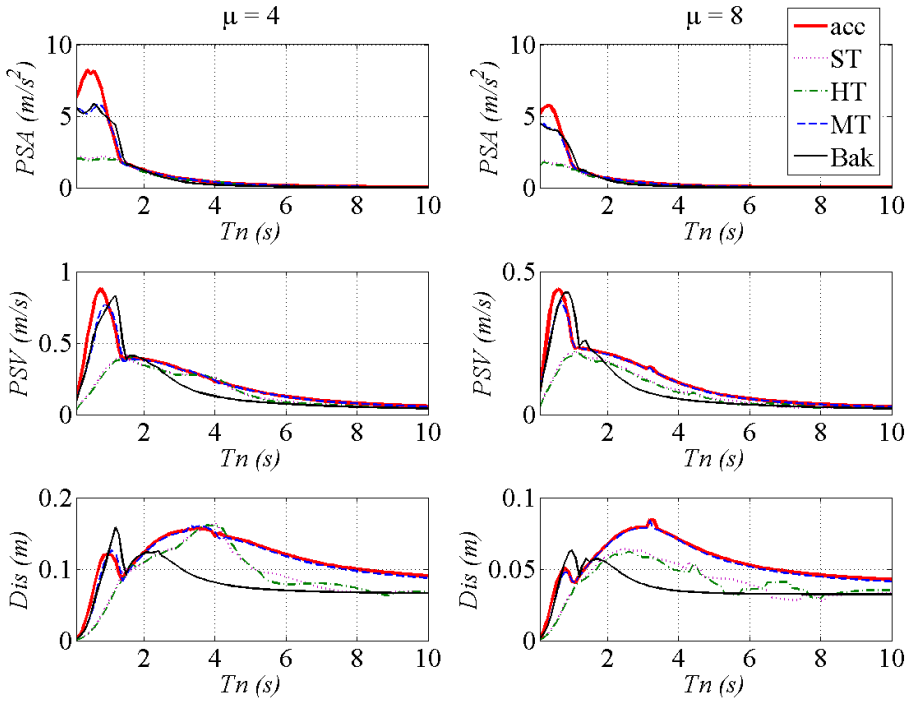

Figure 6.28 1994 Northridge spectral responses for constant ductility factor μ = 4 and μ = 8

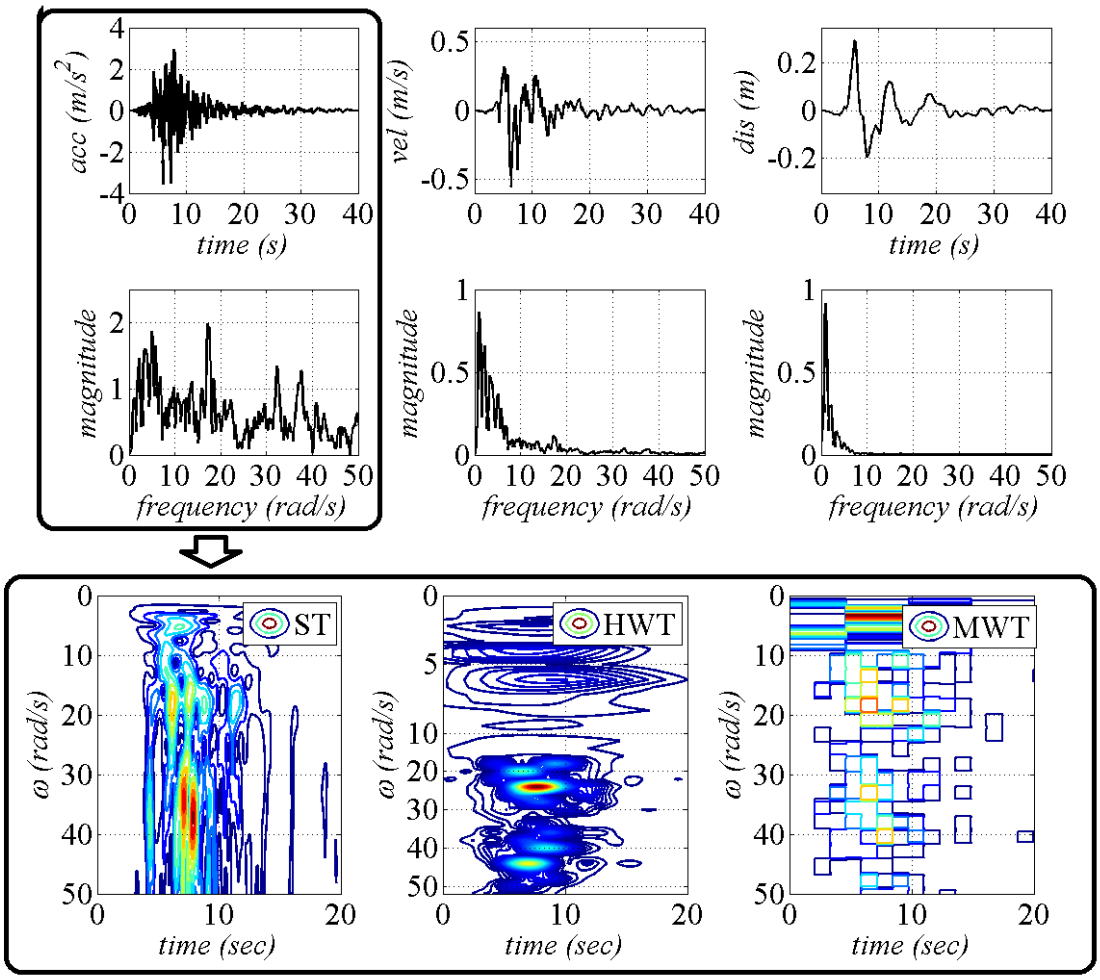

Figure 6.29 1989 Loma Prieta (Saratoga): acceleration, velocity and displacement traces time domain and frequency domain representations.

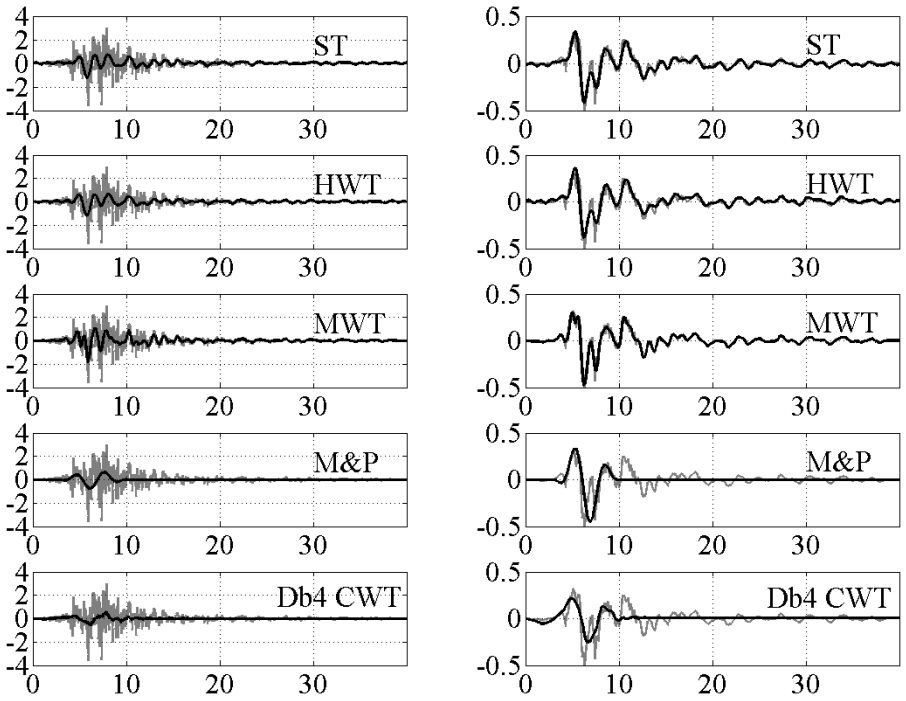

Figure 6.30 1989 Loma Prieta: reconstructed pulses (acceleration and velocity time-histories)

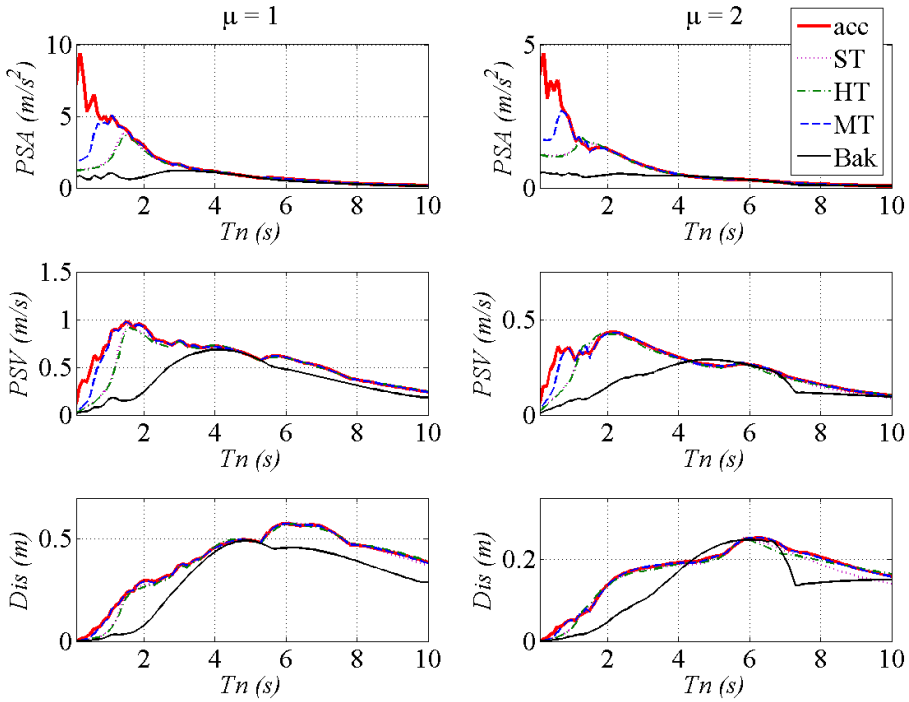

Figure 6.31 1989 Loma Prieta: spectral responses for constant ductility factor μ = 1 (elastic) and μ = 2

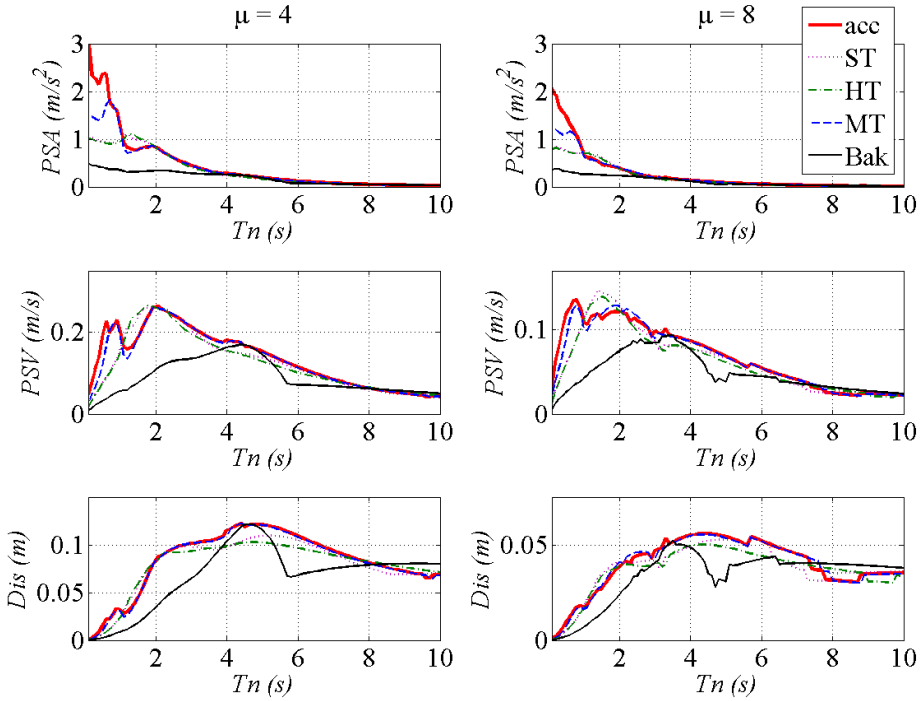

Figure 6.32 1989 Loma Prieta spectral responses for constant ductility factor μ = 4 and μ = 8

CHAPTER 7 : CONCLUDING REMARKS

The present dissertation focuses on the stochastic representation of pulse-like ground motions and on their characterization by means of joint time-frequency representation techniques. The following sections discuss the contributions of this work, the limitations and potential paths for future research.

7.1. SUMMARY OF CONTRIBUTIONS

Pulse-like ground motions attracted the research community interest over the past decade due to the significant demands they pose on a wide range of structures, which may lead to extensive damage in the affected areas. The research work carried herein investigated two aspects of this topic, namely the stochastic modelling of pulse-like ground motion records and the performance of several techniques employed for seismic records' analysis. Consequently, a versatile phenomenological non-stationary stochastic model has been proposed for the representation of pulse-like ground motions. The model was further employed for investigating the potential of several time-frequency analysis techniques to capture the underlying features of the corresponding accelerograms.

Basic signal processing concepts were firstly reviewed, followed by a discussion of several time-frequency analysis methods. Three wavelet-based decomposing techniques, namely the S-transform, the generalized harmonic wavelet transform and the Meyer wavelet packets transform, together with the empirical mode decomposition method were considered for the purposes of this research. Subsequently, the topic of pulse-like ground motions was introduced. The main features of the records and the

structural impact particularities which differentiate them from the ordinary (pulse-free) ones were identified. The physical conditions which facilitate the occurrence of such seismic ground motions were presented, followed by the parameters used for characterizing the specific pulses and the simplified models currently used for identification or, alternatively, for simulation. Finally, the methodologies for pulse-extraction, together with the procedures for the simulation of artificial PLGMs were identified. Based on this review, the following aspects concerning PLGMs records were identified and addressed in this work: the stochastic modelling and the performance of different characterization methods employed.

Considering the amount of uncertainty coming from the limited amount of information available (less than 100 earthquake records classified as being of pulse-type), together with the open debate on the approach to be used for identification and extraction of the characteristic pulses, the stochastic modelling represents an important aspect of the pulse-like ground motions study. The approach commonly used for modelling pulse-like ground motions employs different representation techniques depending on the frequency content considered. Consequently, the low-frequency content in the accelerograms is generally modelled through deterministic velocity models which are differentiated in order to obtain the corresponding acceleration traces. Certain researchers have taken into account the uncertainty in the pulse properties by establishing probability distributions determined for various collections of records for the parameters of the model utilized. Regarding the higher frequency content of the acceleration trace, this is represented by using either seismological-based or record-based stochastic acceleration models. The artificial pulse-like ground motions are obtained by superposing these two representations.

In this work an alternative non-separable non-stationary stochastic model has been developed, which employs a similar approach for modelling the entire frequency

content of the accelerograms. This consists in parametrically defining pulse-like processes as a superposition of several uncorrelated amplitude modulated processes, each modelling the energy distribution in specific regions of the time-frequency plane. In its simplest form, the proposed model can generate pulse-like ground motion processes by considering only two separable uncorrelated processes, one for the low-frequency content and a second one for the higher frequency content. The proposed modelling approach was motivated by the intuitive signal analysis interpretation that pulses represent richer-than-usual low-frequency content of the accelerograms. This observation was made after analysing several records using time-frequency representation techniques.

Two alternative shapes were proposed for modelling the spectral low-frequency content and a time-varying envelope for accounting for its evolutionary behaviour. All the functions were defined using parameters with clear physical meaning. The first shape is a raised cosine, a smooth function suitable for simulation purposes - for example Monte Carlo type of analyses for risk assessment purposes (e.g. Taflanidis & Jia, 2011; Taflanidis, 2011). The second one is a box-like function; due to its simplicity this may be more appropriate in conjunction with random vibrations applications, such as statistical linearization methods (Spanos & Kougioumtzoglou, 2012). The higher frequency content can accommodate representations with various levels of complexity (e.g. Conte & Peng, 1997; Boore, 2003; Yamamoto & Baker, 2013). Herein the uniformly modulated Clough-Penzien spectrum was used for this purpose. Synthetic accelerograms can be readily obtained as realisations of the thus defined pulse-like ground motion model by employing sample generation techniques appropriate for stationary processes.

The capability of the model to realistically capture salient attributes of recorded pulse-like ground motions has been investigated. Consequently, the proposed low-frequency

models have been calibrated against a database extracted pulses reported in the literature. The overall performance in simulating pulses extracted from recorded time-histories with various characteristics, together with their structural impact was deemed satisfactory. Accordingly, regression analyses were performed in order to obtain indicative relationships for defining the pulse model parameters considering various seismic scenarios. In addition, the potential of the models to account for rich low-frequency content (i.e. more than one pulse) was highlighted. Further, the pulse-like ground motion model was calibrated against a field recorded accelerogram. Numerical results indicated that the proposed model leads to similar elastic and inelastic structural responses in average, while accounting for the variability in the characteristics of the ground motion. The applicability of the model for generating code compatible pulse-like accelerograms was also demonstrated by employing an EC8 compatible representation for the higher frequency content. It is believed therefore that the proposed model might represent an alternative to representing the seismic action by means of PLGM spectrum matched accelerograms, which is an issue of interest in earthquake engineering (NEHRP Consultants Joint Venture, 2011).

Once the reliability of the model to represent pulse-like processes and generate artificial time-histories was validated, this was employed to benchmark the potential of various signal processing techniques used in the literature for seismic signal analyses. Their performance in characterizing accelerograms (rather than velocity time-histories) is assessed considering several artificial pulse-like ground motion processes representative for the database of records considered in this work. The argument behind this idea was to demonstrate that the necessary information about PLGMs (i.e. the pulse) can be readily obtained from accelerograms. This is useful since accelerograms are used for processing structural responses; furthermore these are the signals recorded during seismic ground motions, thus less bias induced by data processing techniques affects the

results (Worden, 1990). However, significant amount of research is carried out on velocity traces due to their narrower frequency bandwidth. The numerical results proved that the performance of the TFR techniques considered herein depends on the scope of the analysis. The S-transform is suitable for preliminary identification of pulse-like accelerograms, requiring minimum subjective intervention; the generalized harmonic wavelet transform or Meyer wavelet packets transform appear more suitable for characterization. For filtering out the low-frequency content specific to pulses from the accelerograms, the S-transform and harmonic wavelet transform offered the most accurate results. The suitability of empirical mode decomposition for the characterization on individual records, rather than for databanks was confirmed.

Lastly, the low-frequency content from three field recorded accelerograms was isolated and the spectral responses were evaluated. While the discussion regarding the distinctive border between high and low frequency content remains open, the numerical results showed that this approach accurately captures the structural responses for the case of flexible buildings.

7.2. LIMITATIONS AND FUTURE WORK

A common critique concerning stochastic models for seismic ground motions is that they are perhaps not always capable to realistically account for the variability encountered in recorded ground motions. It has been shown that the model proposed herein allows for variability in the characteristics of the ground motion; however the extent at which this is realistic is still difficult to establish given the amount of data available. Nevertheless, studies on stochastic models showed that by randomizing the model parameters, the variability in the parameters' characteristics can be increased (e.g. Taflanidis & Jia, 2011; Vetter & Taflanidis, 2014). Future studies in this direction are expected to overcome possible limitations in this sense of the developed pulse models.

The potential of the proposed models to lead to realistic structural responses has been investigated from the perspective of the response spectra. It should be kept in mind this offers a limited representation of the structural behaviour (based on the peak displacement) (Huang et al, 2001, NEHRP Consultants Joint Venture, 2011). The reliability of the information provided through the response spectra is questioned by some researchers, since a good representation in the frequency domain (e.g. through the response spectrum) requires a relatively uniform spread of energy over the time (Sommerville 2000, 2002; Bray & Rodriguez-Marek, 2004). At the same time the response spectrum remains a prevalent means to evaluate structural behaviour in the field, due to its simplicity, and thus is deemed satisfactory in the context of this research work.

A further extension of this work could explore the use of the proposed model in conjunction with techniques like statistical linearization in order to approximate the non-linear structural response under pulse-like excitations. Regarding the pulse extraction, further investigations can be done in establishing natural-period-dependant definitions for separating the accelerogram content in low and high frequency.

Aiming to bypass limitations caused by epistemic uncertainties in pulse-like ground motions related research, but also to account for the inherent variability in their properties, an effective fully stochastic model was developed throughout this work which allowed for the investigation of time-frequency techniques potential to characterize the corresponding accelerograms.

APPENDIX A: INTEGRATION AND BASELINE CORRECTIONS VIA FILTERING

A.1. INTEGRATION AS A LOW-PASS FILTERING OPERATION

Consider the acceleration signal *a(t)* having a total duration *T* and its Fourier domain representation $\hat{F}_{acc}(\omega)$. The relationship between the two representations is given by the equation

$$a(t) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \hat{F}_{acc}(\omega) e^{i\omega t} d\omega \tag{A.1}$$

The velocity trace can be obtained from the accelerogram by integration, i.e.

$$v(t) = \int_0^T a(t)\, dt \tag{A.2}$$

This can be also written as

$$\begin{aligned} v(t) &= \int_0^T \left(\frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \hat{F}_{acc}(\omega) e^{i\omega t} d\omega \right) dt = \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \left(\int_0^T e^{i\omega t} dt \right) \hat{F}_{acc}(\omega)\, d\omega = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{1}{i\omega} e^{i\omega t} \hat{F}_{acc}(\omega)\, d\omega \end{aligned} \tag{A.3}$$

which can be rearranged as

$$v(t) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{\hat{F}_{acc}(\omega)}{i\omega} e^{i\omega t} d\omega \tag{A.4}$$

On the other hand, analogous to Eq. (A.1), the relationship between the time domain representation *v(t)* and Fourier domain representations $\hat{F}_{vel}(\omega)$ of the velocity is:

$$v(t) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \hat{F}_{vel}(\omega) e^{i\omega t} d\omega \tag{A.5}$$

From Eq. (A.4) and (A.5) it can be observed that the integration in the time domain is similar to division with $i\omega$ in the frequency domain (Pallás-Areny & Webster, 1999), i.e.:

$$\hat{F}_{vel}(\omega) = \frac{\hat{F}_{acc}(\omega)}{i\omega} \tag{A.6}$$

And consequently the frequency domain representation of the displacement trace will be

$$\hat{F}_{dis}(\omega) = \frac{\hat{F}_{vel}(\omega)}{i\omega} = \frac{\hat{F}_{acc}(\omega)}{(i\omega)^2} \tag{A.7}$$

The multiplication of the frequency domain representation with the factor 1/iω can be interpreted as a filtering of the signal with a filter characterized by the transfer function *H(z)* given bellow (see also Worden, 1990; Pallás-Areny & Webster, 1999).

$$H(z) = \frac{1}{z}, \qquad where \;\; z = i\omega \tag{A.8}$$

The magnitude and the phase of the filter *H(z)*, known as integrator in the signal processing field, are plotted in Figure A.1. Note that *H(z)* diminishes the amplitude of the frequency components as their frequency increases and also performs a -90 deg phase-shift. Consequently, the passage from acceleration to velocity and further to

displacement can be interpreted as a successive low-pass filtering of the initial input signal (i.e. the accelerogram).

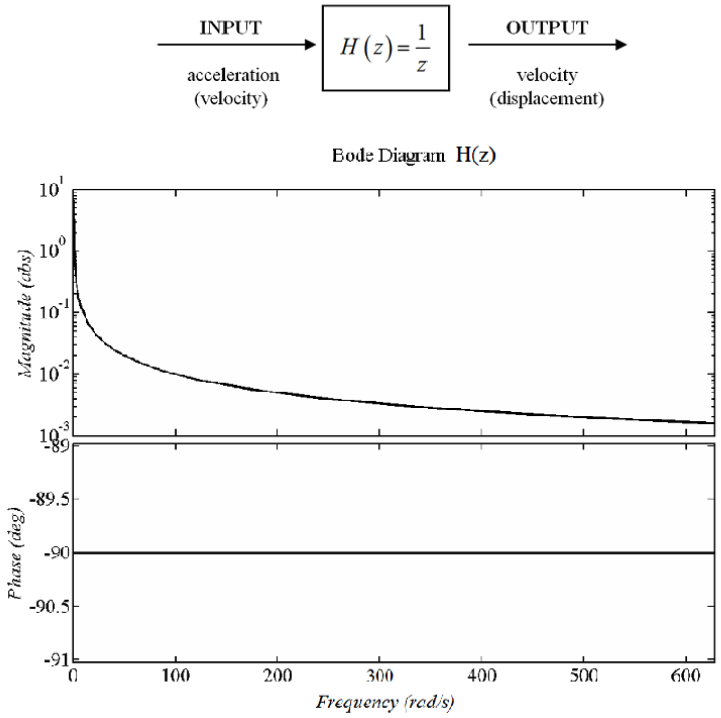

Figure A.1 Frequency domain interpretation of the time-domain integration

A.2. THE BUTTERWORTH FILTER FOR BASELINE CORRECTION

The Butterworth filter is used in the present work for performing the baseline correction of the simulated accelerograms. The filter was introduced in 1930 by St. Butterworth, in an attempt to reduce as much as possible the rippling and the roll-off of the filter's transfer function (obtain a transfer function as flat as possible – see Figure A. 2).

The frequency response function of an order n prototype Butterworth filter characterized by the desired cut-off frequency ω_c is given by the formula (Pallás-Areny & Webster, 1999)

$$H(i\omega)=\frac{1}{1+\left(i\omega/\omega_c\right)^n} \tag{A.9}$$

For the purpose of baseline correcting accelerograms, this filter is modified into a high-pass filter by performing the transformation $i\omega/\omega_c^* \rightarrow \omega_c/i\omega$, where $\omega_c{}^*$ represents the actual cut-off frequency.

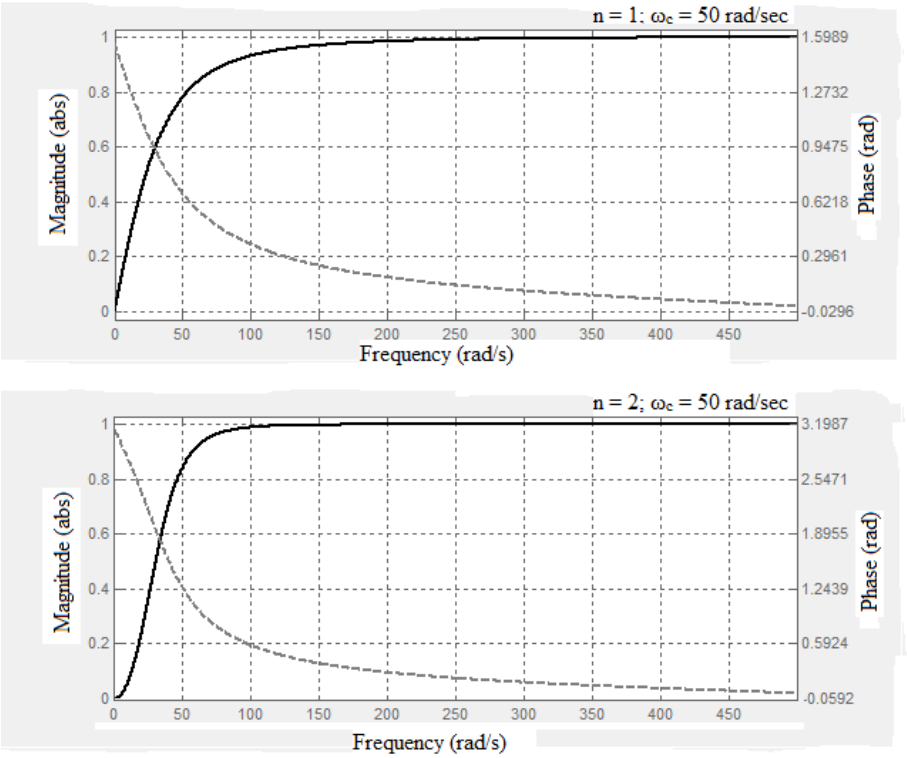

Figure A. 2 High-pass Butterworth filter: cut-off frequency ω_c = 50rad/s and orders n = 1 and 2

APPENDIX B: APPROACHES FOR EARTHQUAKE MODELLING

A basic approach for modelling seismic ground motions utilizes amplitude modulated stationary stochastic processes for this purpose (Shinozuka & Deodatis, 1988; Grigoriu, 1995; Boore, 2003; Giaralis & Spanos, 2009). Consider a stationary process $f_{st}(t)$, whose frequency content is characterized by the power spectrum density function $S_{fst}(\omega)$ and an envelope $a(t)$ with time evolving intensity. A non-stationary random process of the separable kind can be defined by the equation

$$f(t) = a(t) f_{st}(t) \tag{B.1}$$

Under the assumption that the envelope function varies slowly enough, the time-dependant power spectrum density of the process $f(t)$ can be estimated by (Priestley,1965; Conte & Peng, 1997; Spanos & Failla, 2004):

$$S_f(t,\omega) = |a(t)|^2 S_{f_{st}}(\omega) \tag{B.2}$$

Note that the frequency content of the above process remains constant in time; only the time-evolving ground motion intensity which is taken into account.

However, the frequency content of field-recorded accelerograms is also evolving over time. This is caused by the fact that seismic waves characterized by different dominant frequencies arrive at different times. Usually higher frequencies dominate the beginning of the ground motion, with lower frequency components prevailing at later times (Wang

& Zhou, 1999). It is observed that the frequency non-stationarity has an impact on the structural behaviour for both linear and non-linear structures (e.g. Conte & Peng, 1997; Rezaeian & Der Kireghian, 2008). Therefore, realistic stochastic models of seismic records should take into account not only the intensity evolution in time, but also the frequency non-stationarity.

One approach for modelling the frequency non-stationarity consists in utilizing two-parameter time-frequency varying envelopes $a(t,\omega)$ to modulate the stationary process, (Liu, 1970; Yeh & Wen, 1990; Liang, Chaudhuri & Shinozuka, 2007) as in

$$f(t) = a(t,\omega)\, f_{st}(t) \tag{B.3}$$

Analogous to the previous case, the characteristic evolutionary power spectral density can be estimated as

$$S_f(t,\omega) = \left|a(t,\omega)\right|^2 S_{f_{st}}(\omega) \tag{B.4}$$

This is a "fully" non-stationary process of the non-separable kind, as is allows for time-evolution of both intensity and frequency content.

A special class of fully non-stationary processes of particular interest to this work is defined by the summation of R uniformly modulated processes, each one characterized by a specific time envelope and power spectrum density (Spanos & Vargas Loli, 1985; Conte & Peng, 1997; Spanos & Failla, 2004):

$$f(t) = \sum_{r=1}^{R} f_r(t) = \sum_{r=1}^{R} a_r(t)\, f_{st}^r(t) \tag{B.5}$$

As long as the constituent processes $f_r(t)$ are uncorrelated and satisfy the slow-varying envelope condition, the total evolutionary power spectral density can be approximated as the superposition of the power spectra corresponding to each component process, i.e.

$$S_f(t,\omega) = \sum_{r=1}^{R} |a_r(t)|^2 \, S_{f_{st}}{}^r(\omega) \tag{B.6}$$

The power spectral density can be obtained by filtering either Gaussian white noise processes or Poisson processes through time-varying or time-invariant filters (Shinozuka & Deodatis, 1988). In Shinozuka & Deodatis (1988) the two alternatives are compared and it is shown that both lead to the same results up to the second moment under certain specifications. Although Poisson processes are better amenable to physical interpretation, their numerical generation is more computationally demanding. Due to this historical reason, filtered Gaussian white noise is nowadays more commonly encountered in seismic related applications.

Regarding the enveloping function, many shapes have been proposed in the literature, including the Heaviside unit step function, triangular or rectangular functions, various types of exponential functions or piecewise defined functions (Shinozuka & Deodatis, 1988; Quek et al., 1990; Jangid, 2004 among others). The studies of Quek et al. (1990) and Jangid (2004) show that the shape of the envelope has a limited influence on the structural response and it is the energy of the ground motion which needs to be appropriately taken into account.

Realisations of the processes $f(t)$ can be synthesized using appropriate sample generation techniques in order to obtain artificial time-histories.

APPENDIX C: SAMPLE GENERATION TECHNIQUES

C.1. THE SPECTRAL REPRESENTATION METHOD

Realizations of stationary zero-mean random processes can be obtained as a summation of cosines with random phases using the spectral representation method (Shinozuka & Deodatis, 1991).

Consider the uni-dimensional – uni-variable zero-mean stationary stochastic process $f_s(t)$, characterized by the following power spectral density $S_{ff}(\omega)$ and the autocorrelation function $R_{ff}(\tau)$

$$\begin{aligned} &E\left[f_s(t)\right]=0 \\ &E\left[f_s(t+\tau)f_s(t)\right]=R_{ff}(\tau) \\ &R_{ff}(\tau)=\int_{-\infty}^{+\infty}S_{ff}(\omega)e^{i\omega t}d\omega \\ &S_{ff}=\frac{1}{2\pi}\int_{-\infty}^{+\infty}R_{ff}(\tau)e^{-i\omega t}d\tau \end{aligned} \tag{C.1}$$

where $E[\cdot]$ represents the expected value.

According to the Cramer and Leadbetter theorem, a uni-dimensional – uni-variable stationary stochastic process $f_s(t)$ with zero-mean and two-sided power spectrum density $S_{ff}(\omega)$, can be expressed as a function of two real and mutually orthogonal processes $u(\omega)$ and $v(\omega)$, with the increments $du(\omega)$ and as in:

$$f_s(t)=\int_0^{+\infty}[\cos(\omega t)du(\omega)+\sin(\omega t)dv(\omega)] \tag{C.2}$$

The processes $u(\omega)$ and $v(\omega)$ and their increments are defined for $\omega \geq 0$ and have the following properties:

$$\begin{aligned} &E\left[u(\omega)\right] = E\left[v(\omega)\right] = 0 \\ &E\left[u^2(\omega)\right] = E\left[v^2(\omega)\right] = 2S_{FF}(\omega) \\ &E\left[du(\omega)\right] = E\left[dv(\omega)\right] = 0 \\ &E\left[du^2(\omega)\right] = E\left[dv^2(\omega)\right] = 2S_{FF}(\omega)d\omega \end{aligned} \tag{C.3}$$

The previous equations are based on the assumption that $f_s(t)$ can be associated with a differentiable power spectrum distribution $S_{FF}(\omega)$ such that

$$S_{ff}(\omega) = \frac{dS_{FF}(\omega)}{d\omega} \tag{C.4}$$

In the discrete frequency domain a frequency ω_{max} is set after which the values of $S_{ff}(\omega)$ are considered negligible. The power spectrum is discretized in N values with a step $\Delta\omega$. As the value chosen for the step $\Delta\omega$ becomes smaller, the number of discrete values $N \rightarrow \infty$ and the process $f_{st}(t)$ is asymptotically Gaussian.

$$\begin{aligned} &\Delta\omega = \frac{\omega_{\max}}{N} \\ &\omega_k = k\Delta\omega, \quad k = 0,1,\ldots,(N-1) \end{aligned} \tag{C.5}$$

In the discrete time domain, the integral **Error! Reference source not found.** becomes a summation:

$$f(t) = \sum_{k=0}^{N-1}\left[\cos(\omega_k t)\,du(\omega_k) + \sin(\omega_k t)\,dv(\omega_k)\right] \tag{C.6}$$

In Shinozuka & Deodatis (1991) the following form is obtained for the process, after a number of manipulations:

$$f_{st}(t) = \sqrt{2}\sum_{k=0}^{N-1} B_n \cos(\omega_k t + \Phi_k)$$

$$B_n = \sqrt{2S_{ff}(\omega_n)\Delta\omega}, \quad n = 1,2,\ldots,N-1 \tag{C.7}$$

where N is the number of samples in the frequency discretization and $\Phi_k \in [0,2\pi]$ are the random independent phase angles of each cosine. In order to ensure the zero-mean of the samples B_0 needs to be set equal to 0, i.e. $S_{ff}(\omega = 0) = 0$. The resulting samples are periodic, with a period of

$$T_0 = \frac{2\pi}{\Delta\omega} \tag{C.8}$$

and the time-step $\Delta t \leq \pi/\omega_{\max}$ in order to avoid aliasing.

The spectral representation method has been used throughout this work for the purpose of generating realizations of the defined processes.

C.2. AUTOREGRESSIVE-MOVING AVERAGE METHOD

Discrete time realisations of stationary random processes can be obtained by means of the autoregressive-moving average (ARMA) algorithm (Spanos & Mignolet, 1990). This consists in filtering band-limited white noise *w[n]* through an ARMA filter as shown in Figure C.1. This type of filter generates every sample at index *n* as a linear combination between the previous *p* outputs and the contribution of the input white noise, according to the following formulation:

$$f[n] = -\sum_{k=1}^{p} c[k]\, f[n-k] + \sum_{l=0}^{q} b[l]\, w[n-l] \tag{C.9}$$

In the previous equation *c[k]* are coefficients applied to the previous outputs and *b[l]* those applied to the input white noise. The samples $f[n]$ are located $\Delta t = \pi / \omega_{\max}$ (sec) apart one from the other to avoid aliasing, where $0 \le |\omega| \le \omega_{\max}$ is the clipped white noise frequency band. The transfer function for the ARMA filter is

$$H_{ARMA}(z) = \frac{\sum_{l=0}^{q} d_l z^{-k}}{1 + \sum_{k=1}^{p} c_k z^{-k}} \tag{C.10}$$

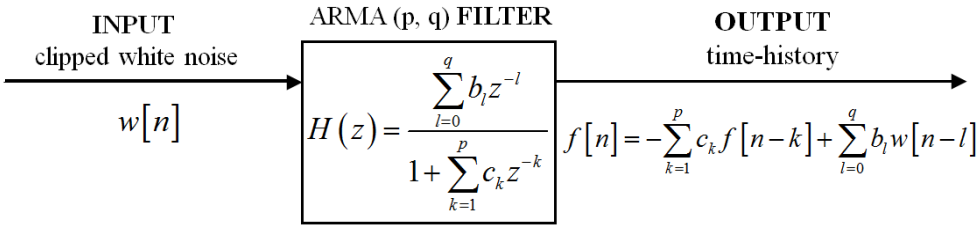

Figure C.1 ARMA method for time-histories simulation

The coefficients of the ARMA filter needed for defining time-histories with a given power spectral density are evaluated by performing the following error minimization:

$$e_{ARMA} \equiv \int_{-\pi/T}^{+\pi/T} \frac{S_{ff}(\omega)}{S_{ARMA}(\omega)}\, d\omega = \min. \tag{C.11}$$

where

$$S_{ARMA}(\omega) = H_{ARMA}\left(e^{i\omega\Delta t}\right) H^{*}{}_{ARMA}\left(e^{i\omega\Delta t}\right) \tag{C.12}$$

The following conditions need to be fulfilled in order to ensure the filter remains stable and causal: the output of the filter has to be finite (stability) and the realisations of the filter for $n < 0$ need to be equal to zero, i.e.:

$$\sum_{n=0}^{\infty} |f[n]| < \infty$$
$$f[n] = 0, \quad for\, n < 0 \tag{C.13}$$

The latter condition is fulfilled by ensuring the values of the white noise are not correlated and the values of the present output depend only on previous input values:

$$\langle w_n, w_{n+k} \rangle = 0, \quad for\, k \neq 0 \quad and \quad \langle f_n, w_{n+k} \rangle = 0, \; for\, k > 0 \tag{C.14}$$

Generally, large numbers of samples are required for obtaining good approximations of the target spectra. From this perspective the spectral representation method can sometimes be computationally expensive; is such cases the auto-regressive moving average method may be more appropriate for sample generation.

APPENDIX D: DATABASE OF PULSE-LIKE RECORDS

Table D.1. Database of pulse-like records considered for the calibration of the low-frequency part of the model

#	Event	Station	Year	M_w	R^* (km)	T_p^{**} (sec)	T_p^{***} (sec)
1	Imperial Valley-06	El Centro Array #3	1979	6.5	12.9	4.0	5.2
2	Imperial Valley-06	El Centro Array #4	1979	6.5	7.1	3.9	4.6
3	Imperial Valley-06	El Centro Array #5	1979	6.5	4.0	3.3	4.0
4	Imperial Valley-06	El Centro Array #6	1979	6.5	1.4	3.5	3.8
5	Imperial Valley-06	El Centro Array #7	1979	6.5	0.6	3.1	4.2
6	Imperial Valley-06	El Centro Array #8	1979	6.5	3.9	4.2	5.4
7	Imperial Valley-06	El Centro Differential Array	1979	6.5	5.1	4.9	5.9
8	Imperial Valley-06	Holtville Post Office	1979	6.5	7.7	4.2	4.8
9	Mammoth Lakes-06	Long Valley Dam (Upr L Abut)	1980	5.9	--	0.9	1.1
10	Irpinia, Italy-01	Sturno	1980	6.9	10.8	2.2	3.1
11	San Fernando	Pacoima Dam (upper left abut)	1971	6.6	1.8	1.2	1.6
12	Westmorland	Parachute Test Site	1981	5.9	16.7	3.1	3.6
13	Coalinga-05	Oil City	1983	5.8	--	0.5	0.7
14	Coalinga-05	Transmitter Hill	1983	5.8	--	0.7	0.9
15	Coalinga-07	Coalinga-14th & Elm (Old CHP)	1983	5.2	--	0.3	0.4
16	Morgan Hill	Coyote Lake Dam (SW Abut)	1984	6.2	0.5	0.7	1.0
17	Morgan Hill	Gilroy Array #6	1984	6.2	9.9	0.9	1.2
18	Taiwan SMART1(40)	SMART1 C00	1986	6.3	--	1.1	1.6
19	Taiwan SMART1(40)	SMART1 M07	1986	6.3	--	1.1	1.6
20	N. Palm Springs	North Palm Springs	1986	6.1	4.0	1.2	1.4
21	San Salvador	Geotech Investig Center	1986	5.8	6.3	0.6	0.9
22	Coyote Lake	Gilroy Array #6	1979	5.7	3.1	0.9	1.2
23	Whittier Narrows-01	Downey - Co Maint Bldg	1987	6.0	20.8	0.8	0.8
24	Whittier Narrows-01	LB - Orange Ave	1987	6.0	24.5	0.7	1.0

#	Event	Station	Year	M_w	R^* (km)	T_p^{**} (sec)	T_p^{***} (sec)
25	Superstition Hills-02	Parachute Test Site	1987	6.5	1.0	2.0	2.3
26	Loma Prieta	Alameda Naval Air Stn Hanger	1989	6.9	71.0	1.7	2.0
27	Loma Prieta	Gilroy Array #2	1989	6.9	11.1	1.3	1.7
28	Loma Prieta	Oakland - Outer Harbor Wharf	1989	6.9	74.3	1.4	1.8
29	Loma Prieta	Saratoga - Aloha Ave	1989	6.9	8.5	4.0	4.5
30	Erzican, Turkey	Erzincan	1992	6.7	4.4	2.1	2.7
31	Cape Mendocino	Petrolia	1992	7.0	8.2	2.2	3.0
32	Landers	Barstow	1992	7.3	34.9	8.0	8.9
33	Imperial Valley-06	Aeropuerto Mexicali	1979	6.5	0.3	1.9	2.4
34	Landers	Lucerne	1992	7.3	2.2	5.3	5.1
35	Landers	Yermo Fire Station	1992	7.3	23.6	6.3	7.5
36	Northridge-01	Jensen Filter Plant	1994	6.7	5.4	2.6	3.5
37	Northridge-01	Jensen Filter Plant Generator	1994	6.7	5.4	2.6	3.5
38	Northridge-01	LA - Wadsworth VA Hospital North	1994	6.7	23.6	2.0	2.4
39	Northridge-01	LA Dam	1994	6.7	5.9	1.3	1.7
40	Northridge-01	Newhall - W Pico Canyon Rd.	1994	6.7	5.5	2.3	2.4
41	Northridge-01	Pacoima Dam (downstr)	1994	6.7	7.0	0.4	0.5
42	Northridge-01	Pacoima Dam (upper left)	1994	6.7	7.0	0.9	0.9
43	Northridge-01	Rinaldi Receiving Sta	1994	6.7	6.5	1.2	1.2
44	Imperial Valley-06	Agrarias	1979	6.5	0.7	1.7	2.3
45	Northridge-01	Sylmar - Converter Sta	1994	6.7	5.4	2.7	3.5
46	Northridge-01	Sylmar - Converter Sta East	1994	6.7	5.2	2.9	3.5
47	Northridge-01	Sylmar - Olive View Med FF	1994	6.7	5.3	2.2	3.1
48	Kobe, Japan	Takarazuka	1995	6.9	0.3	1.1	1.4
49	Kobe, Japan	Takatori	1995	6.9	1.5	1.5	1.6
50	Kocaeli, Turkey	Gebze	1999	7.5	10.9	4.7	5.9
51	Chi-Chi, Taiwan	CHY006	1999	7.6	9.8	2.0	2.6
52	Imperial Valley-06	CHY035	1999	7.6	12.7	1.3	1.4
53	Imperial Valley-06	CHY101	1999	7.6	10.0	4.3	4.8
54	Imperial Valley-06	TAP003	1999	7.6	102.4	2.6	3.4

#	Event	Station	Year	M_w	R^* (km)	T_p^{**} (sec)	T_p^{***} (sec)
55	Imperial Valley-06	Brawley Airport	1979	6.5	10.4	3.8	4.0
56	Imperial Valley-06	TCU029	1999	7.6	28.1	4.7	6.4
57	Imperial Valley-06	TCU031	1999	7.6	30.2	4.5	6.2
58	Imperial Valley-06	TCU034	1999	7.6	35.7	6.4	8.6
59	Imperial Valley-06	TCU036	1999	7.6	19.8	5.6	5.4
60	Mammoth Lakes-07	TCU038	1999	7.6	25.4	6.4	7.0
61	Irpinia, Italy-02	TCU040	1999	7.6	22.1	5.6	6.3
62	San Fernando	TCU042	1999	7.6	26.3	6.4	9.1
63	Westmorland	TCU046	1999	7.6	16.7	6.5	8.6
64	Coalinga-05	TCU049	1999	7.6	3.8	9.0	11.8
65	Coalinga-05	TCU053	1999	7.6	6.0	9.0	12.9
66	Coalinga-07	EC County Center FF	1979	6.5	7.3	4.0	4.5
67	Morgan Hill	TCU054	1999	7.6	5.3	8.2	10.5
68	Morgan Hill	TCU056	1999	7.6	10.5	10.0	12.9
69	Taiwan SMART1(40)	TCU060	1999	7.6	8.5	11.2	12.0
70	Taiwan SMART1(40)	TCU065	1999	7.6	0.6	4.3	5.7
71	N. Palm Springs	TCU068	1999	7.6	0.3	9.0	12.2
72	San Salvador	TCU075	1999	7.6	0.9	4.3	5.1
73	Coyote Lake	TCU076	1999	7.6	2.8	3.1	4.0
74	Whittier Narrows-01	TCU082	1999	7.6	5.2	7.5	9.2
75	Whittier Narrows-01	TCU087	1999	7.6	7.0	8.2	9.0
76	Superstition Hills-03	TCU098	1999	7.6	47.7	6.4	7.5
77	Loma Prieta	EC Meloland Overpass FF	1979	6.5	0.1	2.5	3.3
78	Loma Prieta	TCU101	1999	7.6	2.1	8.2	10.0
79	Loma Prieta	TCU102	1999	7.6	1.5	6.9	9.7
80	Loma Prieta	TCU103	1999	7.6	6.1	7.5	8.3
81	Erzican, Turkey	TCU104	1999	7.6	12.9	10.0	12.0
82	Cape Mendocino	TCU128	1999	7.6	13.2	6.4	9.0
83	Landers	TCU136	1999	7.6	8.3	8.2	10.3
84	Imperial Valley-07	Jiashi	1997	6.1	--	1.0	1.3
85	Landers	Napa Fire Station #3	2000	5.0	--	0.6	0.7

#	Event	Station	Year	M_w	R^* (km)	T_p^{**} (sec)	T_p^{***} (sec)
86	Landers	CHY024	1999	6.2	19.7	2.4	3.2
87	Northridge-01	CHY080	1999	6.2	22.4	1.0	1.4
88	Northridge-01	El Centro Array #10	1979	6.5	6.2	4.1	4.5
89	Northridge-01	TCU076	1999	6.2	14.7	0.8	0.9
90	Northridge-01	CHY101	1999	6.3	36.0	2.1	2.8
91	Northridge-01	El Centro Array #11	1979	6.5	12.5	5.6	7.4

* *Closest distance from the ruptured area to the recording site (when available)*

** *Pulse period estimated in Chapter 5.*

*** *Pulse period estimated by Baker (2007)*

REFERENCES

Alavi, B. & Krawinkler, H. (2000). Consideration of near-fault ground motion effects in seismic design. *Proceedings of the 12th World Conference on Earthquake Engineering.*

Alavi, B. & Krawinkler, H. (2001). Effects of near-fault ground motions on frame structures. *John A. Blume Earthquake Engineering Center, Stanford University, Stanford, California*.

Alavi, B. & Krawinkler, H. (2004). Behavior of moment-resisting frame structures subjected to near-fault ground motions. *Earthquake Engineering & Structural Dynamics*, 33(6), 687-706.

Almufti, I., Motamed, R., Grant, D. N., & Willford, M. (2013). Incorporation of Velocity Pulses in Design Ground Motions for Response History Analysis Using a Probabilistic Framework. *Earthquake Spectra* (preprint).

Askari, R. & Siahkoohi, H.R. (2008). Ground roll attenuation using the S and x-f-k transforms. *Geophysical Prospecting*, 56(1), pp.105-114.

Baker, J.W. (2007). Quantitative classification of near-fault ground motions using wavelet analysis. *Bulletin of the Seismological Society of America*, 97(5), pp.1486-1501.

Bogdanoff, J. L., Goldberg, J. E. & Bernard, M. C. (1961). Response of a simple structure to a random earthquake-type disturbance. *Bulletin of the Seismological Society of America*, 51(2), pp. 293-310.

Boore, D.M. (2003). Simulation of ground motion using the stochastic method. *Pure and Applied Geophysics*, 160(3-4), pp.635-676.

Boore, D. M., & Akkar, S. (2003). Effect of causal and acausal filters on elastic and inelastic response spectra. *Earthquake Engineering & Structural Dynamics*, 32(11), pp. 1729-1748.

Boore, D. M. & Bommer, J. J. (2005). Processing of strong-motion accelerograms: needs, options and consequences. *Soil Dynamics and Earthquake Engineering*, 25(2), pp. 93-115.

Bray, J.D. & Rodriguez-Marek, A. (2004). Characterization of forward-directivity ground motions in the near-fault region. *Soil Dynamics and Earthquake Engineering*, 24(11), pp. 815-828.

Cacciola, P. & Deodatis, G. (2011). A method for generating fully non-stationary and spectrum-compatible ground motion vector processes. *Soil Dynamics and Earthquake Engineering*, 31(3), pp. 351-360.

CEN (2004). Eurocode 8: Design of Structures for Earthquake Resistance-Part 1: General Rules, Seismic Actions and Rules for Buildings. *EN 1998-1: 2004, Comité Européen de Normalisation, Brussels*.

Chioccarelli, E. & Iervolino, I. (2013). Near-source seismic hazard and design scenarios. *Earthquake Engineering & Structural Dynamics*, 42(4), 603-622.

Clough, R.W. & Penzien, J. (1993). *Dynamics of structures*. 2nd ed. New York: McGraw-Hill.

Cohen, L. (1995). *Time-frequency analysis: theory and applications*. Ed. Englewood Cliffs, NJ: Prentice-Hall.

Conte, J.P. & Peng, B.F. (1997). Fully nonstationary analytical earthquake ground-motion model. *Journal of Engineering Mechanics*, 123(1), pp.15-24.

Dabaghi, M. & Der Kireghian, A. (2011). Stochastic model and simulation of near fault ground motions for specified earthquake source and site characteristics' (Report). *California Department of Transportation*.

Dabaghi, M., Rezaeian, S. & Der Kiureghian, A. (2011). Stochastic simulation of near-fault ground motions for specified earthquake and site characteristics. *Proceedings of the 11th International Conference on Applications of Statistics and Probability in Civil Engineering, Zurich, Switzerland*, pp. 2498-2505.

Daubechies, I. (1992). *Ten lectures on wavelets* (vol.61). Society for Industrial and Applied Mathematics, Philadelphia.

Dätig, M.& Schlurmann, T. (2004). Performance and limitations of the Hilbert-Huang transformation (HHT) with an application to irregular water waves. *Ocean Engineering*, 31(14-15), pp.1783-1834.

Dickinson, B.W. & Gavin, H.P. (2011). Parametric statistical generalization of uniform-hazard earthquake ground motions. *Journal of structural engineering* 137(3), pp.410-422.

Dugundji, J. (1958). Envelopes and pre-envelopes of real wave-forms. *Information Theory, IRE Transactions*, 4, pp.53-57.

Farge, M. (1992). Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics*, 24(1), pp. 395-457.

Farid Ghahari, S., Jahankhah, H. & Ghannad, M. A. (2010). Study on elastic response of structures to near-fault ground motions through record decomposition. *Soil dynamics and earthquake engineering*, 30(7), pp. 536-546.

Flandrin, P., Rilling, G. & Gonc□alvés, P. (2004). Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters,* 11 (2 Part I) pp.112-114.

Fu, Q. & Menun, C. (2004). Seismic-environment-based simulation of near-fault ground motions. *13th World Conference on Earthquake Engineering, Vancouver, B.C., Canada,* no. 322, 15pp.

Gavin, H.P. & Dickinson, B.W. (2011). Generation of uniform-hazard earthquake ground motions. *Journal of Structural Engineering,* 137(3), pp. 423-432.

Giaralis, A. & Lungu, A. (2012). Assessment of wavelet-based representation techniques for the characterization of stochastic processes modelling pulse-like strong ground motions. *Proceedings of the 6th International ASRANet Conference for Integrating Structural Analysis, Risk and Reliability, London.*

Giaralis, A. & Spanos, P.D. (2009). Wavelet-based response spectrum compatible synthesis of accelerograms-Eurocode application (EC8). *Soil Dynamics and Earthquake Engineering,* 29(1), pp. 219-235.

Giaralis, A. & Spanos, P.D. (2012). Derivation of response spectrum compatible non-stationary stochastic processes relying on Monte Carlo-based peak factor estimation. *Earthquake and Structures,* 3(3), pp. 581-609.

Grigoriu, M. (1995). Probabilistic models and simulation methods for seismic ground acceleration. *Meccanica,* 30(1), pp. 105-124.

Guo, W., Tse, P.W. & Djordjevich, A. (2012). Faulty bearing signal recovery from large noise using a hybrid method based on spectral kurtosis and ensemble empirical mode decomposition. *Measurement: Journal of the International Measurement Confederation,* 45(5), pp. 1308-1322.

He, W. & Agrawal, A.K. (2008). Analytical model of ground motion pulses for the design and assessment of seismic protective systems. *Journal of Structural Engineering,* 134(7), pp. 1177-1188.

Huang, G. & Chen, X. (2009). Wavelets-based estimation of multivariate evolutionary spectra and its application to nonstationary downburst winds. *Engineering Structures,* 31(4), pp. 976-989.

Huang, N.E., Chern, C.C., Huang, K., Salvino, L.W., Long, S.R. & Fan, K.L. (2001). A new spectral representation of earthquake data: Hilbert spectral analysis of station TCU129, Chi-Chi, Taiwan, 21 September 1999. *Bulletin of the Seismological Society of America,* 91(5), pp. 1310-1338.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Snin, H.H., Zheng, Q., Yen, N., Tung, C.C. & Liu, H.H. (1998). The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 454(1971), pp. 903-995.

Huang, N.E., Wu, M.C., Long, S.R., Shen, S.S.P., Qu, W., Gloersen, P. & Fan, K.L. (2003). A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 459(2037), pp. 2317-2345.

Iervolino, I. & Cornell, C.A. (2008). Probability of occurrence of velocity pulses in near-source ground motions. *Bulletin of the Seismological Society of America,* 98(5), pp. 2262-2277.

Jangid, R.S. (2004). Response of SDOF system to non-stationary earthquake excitation. *Earthquake Engineering & Structural Dynamics,* 33(15), pp. 1417-1428.

Kanai, K. (1957). Semi-empirical formula for the seismic characteristics of the ground. *University of Tokyo, Bulletin of the Earthquake Research Institute*, 35, pp. 309–325.

Karabalis, D., Cokkinides, G.Rizos, D. & Mulliken, J. (2000). Simulation of earthquake ground motions by a deterministic approach. *Advances in Engineering Software,* 31(5), pp. 329-338.

Kijewski-Correa, T. & Kareem, A. (2006). Efficacy of Hilbert and wavelet transforms for time-frequency analysis. *Journal of Engineering Mechanics,* 132(10), pp. 1037-1049.

Lai, S.P. (1982). Statistical characterization of strong ground motions using power spectral density function. *Bulletin of the Seismological Society of America*, 72(1), pp. 259-274.

Liang, J., Chaudhuri, S. R. & Shinozuka, M. (2007). Simulation of nonstationary stochastic processes by spectral representation. *Journal of engineering mechanics*, 133(6), pp. 616-627.

Liu, S. C. (1970). Evolutionary power spectral density of strong-motion earthquakes. *Bulletin of the Seismological Society of America*, 60(3), pp. 891-900.

Loh, C., Wu, T. & Huang, N.E. (2001). Application of the empirical mode decomposition-Hilbert spectrum method to identify near-fault ground-motion characteristics and structural responses. *Bulletin of the Seismological Society of America,* 91(5), pp. 1339-1357.

Lungu, A. & Giaralis, A. (2013). A non-separable stochastic model for pulse-like ground motions. *Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infrastructures - Proceedings of the 11th International Conference on Structural Safety and Reliability, ICOSSAR 2013 New York, New York,* pp. 1017-1024.

Makris, N. & Chang, S.P. (2000a). Effect of viscous, viscoplastic and friction damping on the response of seismic isolated structures. *Earthquake Engineering & Structural Dynamics*, 29(1), pp. 85-107.

Makris, N. & Black, C.J. (2004a). Dimensional analysis of rigid-plastic and elastoplastic structures under pulse-type excitations. *Journal of Engineering Mechanics*, 130(9), pp. 1006-1018.

Makris, N. & Black, C.J. (2004b). Dimensional analysis of bilinear oscillators under pulse-type excitations. *Journal of Engineering Mechanics*, 130(9), pp. 1019-1031.

Makris, N. & Black, C.J. (2004c). Evaluation of peak ground velocity as a "good" intensity measure for near-source ground motions. *Journal of Engineering Mechanics*, 130(9), pp. 1032-1044.

Makris, N. & Chang, S.P. (2000b). Response of damped oscillators to cycloidal pulses. *Journal of Engineering Mechanics*, 126(2), pp. 123-131.

Malhotra, P.K. (1999). Response of buildings to near-field pulse-like ground motions. *Earthquake Engineering & Structural Dynamics*, 28(11), pp. 1309-1326.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7), pp. 674-693.

Mallat, S.G. (2009). *A wavelet tour of signal processing: the sparse way*. Amsterdam; Boston: Elsevier /Academic Press.

Mann, S. & Haykin, S. (1999) 'The chirplet transform: physical considerations', *Signal Processing, IEEE Transactions*, 43(11), pp. 2745-2761.

Mavroeidis, G.P., Dong, G. & Papageorgiou, A.S. (2004). Near-fault ground motions, and the response of elastic and inelastic single-degree-of-freedom (SDOF) systems. *Earthquake Engineering and Structural Dynamics*, 33(9), pp. 1023-1049.

Mavroeidis, G.P. & Papageorgiou, A.S. (2003). A mathematical representation of near-fault ground motions. *Bulletin of the Seismological Society of America*, 93(3), pp. 1099-1131.

McFadden, P.D., Cook, J.G. & Forster, L.M. (1999). Decomposition of gear vibration signals by the generalized S transform. *Mechanical Systems and Signal Processing*, 13(5), pp. 691-708.

Menun, C. & Fu, Q. (2002). An analytical model for near-fault ground motions and the response of SDOF systems. *Seventh US National Conference on Earthquake Engineering, Boston, Massachusetts*.

Mollaioli, F. & Bosi, A. (2012). Wavelet analysis for the characterization of forward-directivity pulse-like ground motions on energy basis. *Meccanica*, 47(1), pp. 203-219.

Moustafa, A. (2010). Discussion of "Analytical model of ground motion pulses for the design and assessment of seismic protective systems" by W.-L. He and A. K. Agrawal. *Journal of Structural Engineering*, 136(2), pp. 229-230.

Moustafa, A. & Takewaki, I. (2010). Deterministic and probabilistic representation of near-field pulse-like ground motion. *Soil Dynamics and Earthquake Engineering,* 30(5), pp. 412-422.

Mukherjee, S. & Gupta, V. K. (2002). Wavelet-based generation of spectrum-compatible time-histories. *Soil Dynamics and Earthquake Engineering*, 22(9), pp. 799-804.

Mukhopadhyay, S. & Gupta, V. K. (2013a). Directivity pulses in near-fault ground motions-I: Identification, extraction and modelling. *Soil Dynamics and Earthquake Engineering,* 50, pp. 1-15.

Mukhopadhyay, S. & Gupta, V. K. (2013b). Directivity pulses in near-fault ground motions-II: Estimation of pulse parameters. *Soil Dynamics and Earthquake Engineering,* 50, pp. 38-52.

NEHRP Consultants Joint Venture (2011). Selecting and scaling earthquake ground motions for performing response-history analyses. *National Institute of Standards and Technology: Redwood City, California*.

Newland, D.E. (1984). *An introduction to random vibrations and spectral analysis.* 2nd edn. Harlow: Longman, New York.

Newland, D.E. (1994). Harmonic and musical wavelets. *Proceedings of the Royal Society A, Mathematical and physical sciences,* 444(1922), pp. 605-620.

Newland, D.E. (1999). Ridge and phase identification in the frequency analysis of transient signals by harmonic wavelets. *Journal of Vibration and Acoustics, Transactions of the ASME,* 121(2), pp. 149-155.

Pallás-Areny, R. & Webster, J.G. (1999). *Analog signal processing*. Chichester; New York: Wiley.

Parolai, S. (2009). Denoising of seismograms using the S transform. *Bulletin of the Seismological Society of America,* 99(1), pp. 226-234.

Pinnegar, C.R. & Eaton, D.W. (2003). Application of the S transform to prestack noise attenuation filtering. *Journal of Geophysical Research: Solid Earth (1978-2012),* 108(B9), pp. 9-1 - 9-10.

Pinnegar, C.R. & Mansinha, L. (2004). Time-local Fourier analysis with a scalable, phase-modulated analyzing function: The S-transform with a complex window. *Signal Processing,* 84(7), pp. 1167-1176.

Priestley, M.B. (1965). Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society. Series B,* pp. 204-237.

Qian, S. (2002). *Introduction to time-frequency and wavelet transforms.* Upper Saddle River, N.J.: Prentice Hall.

Quek, S.T., Teo, Y.P. & Balendra, T. (1990). Non-stationary structural response with evolutionary spectra using seismological input model. *Earthquake Engineering & Structural Dynamics*, 19(2), pp. 275-288.

Rezaeian, S. & Der Kiureghian, A. (2008). A stochastic ground motion model with separable temporal and spectral nonstationarities. *Earthquake Engineering and Structural Dynamics*, 37(13), pp. 1565-1584.

Rezaeian, S. & Der Kiureghian, A. (2010). Stochastic modelling and simulation of ground motions for performance-based earthquake engineering. *Pacific Earthquake Engineering Research Center, University of California, Berkeley*.

Rilling, G., Flandrin, P. & Goncalves, P. (2003). On empirical mode decomposition and its algorithms. *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP*, pp. 8.

Safak, E. & Boore, D. M. (1988). On low-frequency errors of uniformly modulated filtered white-noise models for ground motions. *Earthquake Engineering & Structural Dynamics*, 16(3), 381-388.

Sehhati, R., Rodriguez-Marek, A., El Gawady, M. & Cofer, W.F. (2011). Effects of near-fault ground motions and equivalent pulses on multi-story structures. *Engineering Structures*, 33(3), pp. 767-779.

Shahi, S.K. & Baker, J.W. (2011). An empirically calibrated framework for including the effects of near-fault directivity in probabilistic seismic hazard analysis. *Bulletin of the Seismological Society of America*, 101(2), pp. 742-755.

Shinozuka, M. & Deodatis, G. (1988). Stochastic process models for earthquake ground motion. *Probabilistic Engineering Mechanics*, 3(3), pp. 114-123.

Shinozuka, M. & Deodatis, G. (1991). Simulation of stochastic processes by spectral representation. *Applied Mechanics Reviews*, 44(4), pp. 191-204.

Somerville, P. (1997). Engineering characteristics of near fault ground motion. *SMIP97 Seminar on Utilization of Strong-Motion Data, Los Angeles, California.*

Somerville, P. (1998). Development of an improved representation of near fault ground motions. *SMIP98 Seminar on Utilization of Strong-Motion Data, Oakland, California*.

Somerville, P. (2000). Magnitude scaling of near fault ground motions. *EOS, Trans. Am. Geophys. Union*, 81.

Somerville, P. (2002). Characterizing near fault ground motion for the design and evaluation of bridges. *Third National Conference and Workshop on Bridges and Highways. Portland, Oregon.*

Somerville, P.G. (2003). Magnitude scaling of the near fault rupture directivity pulse. *Physics of the Earth and Planetary Interiors*, 137(1), pp. 201-212.

Spanos, P.D. & Failla, G. (2004). Evolutionary spectra estimation using wavelets. *Journal of Engineering Mechanics,* 130(8), pp. 952-960.

Spanos, P. D. & Failla, G. (2005). Wavelets: Theoretical concepts and vibrations related applications. *The Shock and vibration digest*, 37(5), pp. 359-375.

Spanos, P. D. & Giaralis, A. (2008). Statistical linearization based estimation of the peak response of nonlinear systems subject to the EC8 design spectrum. *Seismic engineering conference commemorating the 1908 Messina and Reggio Calabria earthquake, Parts 1 and 2*, 1020, pp. 1236-1244.

Spanos, P.D. & Giaralis, A. (2013). Third-order statistical linearization-based approach to derive equivalent linear properties of bilinear hysteretic systems for seismic response spectrum analysis. *Structural Safety,* 44, pp. 59-69.

Spanos, P.D., Giaralis, A. & Li. J (2009). Synthesis of accelerograms compatible with the Chinese GB 50011-2001 design spectrum via harmonic wavelets: artificial and historic records. *Earthquake Engineering and Engineering Vibration,* 8(2), pp. 189-206.

Spanos, P.D., Giaralis, A. & Politis, N.P. (2007). Time-frequency representation of earthquake accelerograms and inelastic structural response records using the adaptive chirplet decomposition and empirical mode decomposition. *Soil Dynamics and Earthquake Engineering,* 27(7), pp. 675-689.

Spanos, P.D. & Kougioumtzoglou, I.A. (2012). Harmonic wavelets based statistical linearization for response evolutionary power spectrum determination. *Probabilistic Engineering Mechanics,* 27(1), pp. 57-68.

Spanos, P. D. & Mignolet, M. P. (1990). Simulation of stationary random processes: two-stage MA to ARMA approach. *Journal of engineering mechanics*, 116(3), pp. 620-641.

Spanos, P. D., Tezcan, J. & Tratskas, P. (2005). Stochastic processes evolutionary spectrum estimation via harmonic wavelets. *Computer Methods in Applied Mechanics and Engineering,* 194*(*12), pp. 1367-1383.

Spanos, P.D. & Vargas Loli, L.M. (1985). Statistical approach to generation of design spectrum compatible earthquake time histories. *Soil Dynamics and Earthquake Engineering,* 4(1), pp. 2-8.

Spanos, P. D. & Zeldin, B. A. (1998). Monte Carlo treatment of random fields: a broad perspective. *Applied Mechanics Reviews*, 51(3), pp. 219-237.

Stafford, P. J., Sgobba, S., & Marano, G. C. (2009). An energy-based envelope function for the stochastic simulation of earthquake accelerograms. *Soil Dynamics and Earthquake Engineering*, 29(7), pp. 1123-1133.

Stockwell, R.G. (1999). S-transform analysis of gravity wave activity from a small scale network of airglow imagers. *PhD thesis, The University of Western Ontario, London, Ontario.*

Stockwell, R.G., Mansinha, L. & Lowe, R.P. (1996). Localization of the complex spectrum: the S transform. *IEEE Transactions on Signal Processing,* 44(4), pp. 993.

Taflanidis, A.A. (2011). Optimal probabilistic design of seismic dampers for the protection of isolated bridges against near-fault seismic excitations. *Engineering Structures,* 33(12), pp. 3496-3508.

Taflanidis, A.A. & Jia, G. (2011). A simulation-based framework for risk assessment and probabilistic sensitivity analysis of base-isolated structures. *Earthquake Engineering and Structural Dynamics,* 40(14), pp. 1629-1651.

Taflanidis, A.A. & Vetter, C. (2011). Seismic risk sensitivity analysis focusing on stochastic ground motion modelling. *Proceedings of the 11th International Conference on Applications of Statistics and Probability in Civil Engineering, Zürich, Switzerland,* pp. 2465-2473.

Tang, Y. & Zhang, J. (2011). Response spectrum-oriented pulse identification and magnitude scaling of forward directivity pulses in near-fault ground motions. *Soil Dynamics and Earthquake Engineering,* 31(1), pp. 59-76.

Teolis, A. (1998). *Computational signal processing with wavelets.* Boston, Mass: Birkhäuser.

Tian, Y. J., Yang, Q. S. & Lu, M. Q. (2007). Simulation method of near-fault pulse-type ground motion. *Acta Seismologica Sinica*, 20, pp. 80-87.

Todorovska, M.I., Meidani, H. & Trifunac, M.D. (2009). Wavelet approximation of earthquake strong ground motion-goodness of fit for a database in terms of predicting nonlinear structural response. *Soil Dynamics and Earthquake Engineering,* 29(4), pp. 742-751.

Torrence, C. & Compo, G.P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society,* 79(1), pp. 61-78.

Tothong, P. & Cornell, C.A. (2008). Structural performance assessment under near-source pulse-like ground motions using advanced ground motion intensity measures. *Earthquake Engineering and Structural Dynamics,* 37(7), pp. 1013-1037.

Travasarou, T., Bray, J. D. & Abrahamson, N. A. (2003). Empirical attenuation relationship for Arias Intensity. *Earthquake Engineering & Structural Dynamics,* 32(7), pp. 1133-1155.

Trifunac, M.D. & Brady, A.G. (1975). A study on the duration of strong earthquake ground motion. *Bulletin of the Seismological Society of America,* 65(3), pp. 581-626.

Vassiliou, M.F. & Makris, N. (2011). Estimating time scales and length scales in pulselike earthquake acceleration records with wavelet analysis. *Bulletin of the Seismological Society of America*, 101(2), pp. 596-618.

Ventosa, S., Simon, C., Schimmed, M., Dañobeitia, J.J. & Mànuel, A. (2008). The S-transform from a wavelet point of view. *IEEE Transactions on Signal Processing*, 56(7I), pp. 2771-2780.

Vetter, C., & Taflanidis, A. A. (2014). Comparison of alternative stochastic ground motion models for seismic risk characterization. *Soil Dynamics and Earthquake Engineering*, 58, pp. 48-65.

Wang, J. & Zhou, J. (1999). Aseismic Designs Based on Artificial Simulations. *IEEE Signal Processing Magazine*, 16(2), pp. 94-99.

Worden, K. (1990). Data processing and experiment design for the restoring force surface method, part I: integration and differentiation of measured time data. *Mechanical Systems and Signal Processing*, 4(4), pp. 295-319.

Xie, L., Xu, L. & Rodriguez-Marek, A. (2005). Representation of near-fault pulse-type ground motions. *Earthquake Engineering and Engineering Vibration*, 4(2), pp. 191-199.

Xu, C., Chen, G. & Xie, J. (2010). Damage detection of offshore platform model using empirical mode decomposition and Wigner-Ville distribution. *Proceeding of the Twentieth International of Offshore and Polar Engineering Conference, Beijing, China ISOPE*, Vol. 4, pp. 644-652.

Xu, Z. & Agrawal, A. (2010). Decomposition and effects of pulse components in near-field ground motions. *Journal of Structural Engineering*, 136(6), pp. 690-699.

Yaghmaei-Sabegh, S. (2010). Detection of pulse-like ground motions based on continues wavelet transform. *Journal of Seismology*, 14 (4), pp. 715-726.

Yamamoto, Y. (2011). Stochastic model for earthquake ground motion using wavelet packets. *PhD thesis, Stanford University, California.*

Yamamoto, Y. & Baker, J.W. (2013). Stochastic model for earthquake ground motion using wavelet packets. *Bulletin of the Seismological Society of America*, 103(6), pp. 3044-3056.

Yeh, C. H. & Wen, Y. K. (1990). Modeling of nonstationary ground motion and analysis of inelastic structural response. *Structural Safety*, 8(1), pp. 281-298.

Yinfeng, D., Yingmin, L., Mingkui, X. & Ming, L. (2008). Analysis of earthquake ground motions using an improved Hilbert-Huang transform. *Soil Dynamics and Earthquake Engineering*, 28(1), pp. 7-19.

Zamora, M. & Riddell, R. (2011). Elastic and inelastic response spectra considering near-fault effects. *Journal of Earthquake Engineering*, 15(5), pp. 775-808.

Zhang, R.R., Ma, S. & Hartzell, S. (2003). Signatures of the seismic source in EMD-based characterization of the 1994 Northridge, California, earthquake recordings. *Bulletin of the Seismological Society of America*, 93 (1), pp. 501-518.

Zhang, Y., Hu, Y., Zhao, F., Liang, J. & Yang, C. (2005). Identification of acceleration pulses in near-fault ground motion using the EMD method. *Earthquake Engineering and Engineering Vibration*, 4(2), pp. 201-212.